

Tue Jan 20 17:43:17 2004

US-10-022-025a-1.rn1

Page 4

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 09:48:47 ; Search time 124 Seconds
(without alignments)
7151.119 Million cell updates/sec

Title: US-10-022-025a-1

Perfect score: 2009
Sequence: 1 aataataataataatgta.....atagagatgagatatt 2009

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/5B_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	420.8	20.9	1611	3	US-08-948-564-7 Sequence 7, Appli
2	200.4	10.0	1799	2	US-08-560-998-3 Sequence 3, Appli
3	123.2	6.1	1704	3	US-08-948-564-17 Sequence 17, Appli
4	101.2	5.0	1708	3	US-08-991-677-1 Sequence 1, Appli
5	98.2	4.9	275	4	US-09-313-294A-2591 Sequence 2591, Ap
6	89	4.4	1788	3	US-08-948-564-9 Sequence 9, Appli
7	88.8	4.4	1893	1	US-08-532-065B-1 Sequence 1, Appli
8	84.6	4.2	2174	3	US-08-606-505B-63 Sequence 63, Appli
9	84.6	4.2	2174	3	US-08-606-505B-63 Sequence 63, Appli
10	83.2	4.1	1761	3	US-08-033-055A-3 Sequence 3, Appli
11	83.2	4.1	1927	3	US-08-606-505B-64 Sequence 64, Appli
12	83.2	4.1	1927	3	US-08-606-505B-64 Sequence 64, Appli
13	81.4	4.1	5156	2	US-09-091-432-3 Sequence 3, Appli
14	81.4	4.1	5156	2	US-09-387-663-3 Sequence 3, Appli
15	79.2	3.9	1722	3	US-09-033-055A-1 Sequence 1, Appli
16	77.2	3.8	1755	1	US-07-912-900-29 Sequence 29, Appli
17	77.2	3.8	1755	1	US-08-285-309-29 Sequence 29, Appli
18	77.2	3.8	1755	1	US-08-502-046-29 Sequence 29, Appli
19	76.4	3.8	1812	1	US-08-313-075A-37 Sequence 37, Appli
20	76.4	3.8	1824	3	US-08-606-505B-1 Sequence 1, Appli
21	76.4	3.8	1824	3	US-08-606-505B-1 Sequence 1, Appli
22	75.6	3.8	1834	4	US-09-615-192A-232 Sequence 232, App
23	75.6	3.7	1835	4	US-09-564-808-5 Sequence 5, Appli
24	75	3.7	1838	2	US-09-091-432-1 Sequence 1, Appli
25	75	3.7	1838	4	US-09-387-663-1 Sequence 1, Appli
26	73	3.6	1620	4	US-09-627-216A-11 Sequence 11, Appli

28	72.8	3.6	543	4	US-09-615-192A-230 Sequence 230, App
29	71.8	3.6	1884	4	US-09-564-808-3 Sequence 3, Appli
30	70.2	3.5	1657	3	US-08-948-564-11 Sequence 11, Appli
31	69.2	3.4	1806	4	US-09-351-229-3 Sequence 3, Appli
32	68.6	3.4	1880	4	US-09-564-808-1 Sequence 1, Appli
33	66.6	3.3	1781	4	US-09-499-302A-1 Sequence 1, Appli
34	65.4	3.3	1812	1	US-07-912-900-28 Sequence 28, Appli
35	65.4	3.3	1812	1	US-08-285-309-28 Sequence 28, Appli
36	65.4	3.3	1812	2	US-08-502-046-28 Sequence 28, Appli
37	64	3.2	1506	3	US-09-158-767-7 Sequence 7, Appli
38	64	3.2	1506	3	US-09-158-767-8 Sequence 8, Appli
39	64	3.2	1506	3	US-09-158-767-9 Sequence 9, Appli
40	64	3.2	2261	3	US-09-158-767-1 Sequence 1, Appli
41	63	3.1	1737	4	US-09-126-420A-1 Sequence 1, Appli
42	63	3.1	1824	3	US-08-948-564-13 Sequence 13, Appli
43	62	3.1	1824	3	US-08-948-564-13 Sequence 13, Appli
44	61.4	3.1	1634	1	US-08-313-075A-55 Sequence 55, Appli
45	61.4	3.1	1644	4	US-09-126-420A-2 Sequence 2, Appli
					US-08-948-564-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-948-564-7
Sequence 7, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminsky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corlin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512ch Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCES/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
FAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..1588
US-08-948-564-7
Query Match
Best Local Similarity 20.9%
Score 420.8; DB 3; Length 1611;
Pred. No. 4.9e-107;
Matches 1029; Conservative 0; Mismatches 487; Indels 256; Gaps 6;

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rnt

Page 2

221 AGGCAACCAACCTGCTTCCTCCCTCCGCGCCGCAATCATCTGCTGCGCATCTT 280
74 ACCCAAGAAACCTGCTGATGGCTCTTTGATCATGGCTCACTCTGCTTAACTACT 133
281 CTCTCTTAATGACCTATCCGCTGACCTGCTTGGGGAATACCTCTTCCGCGCTTA 340
134 TTTATTAATGCTGACACCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 186
341 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 400
187 CT 235
401 GGAAGCATCTCATCTCATCTCATCTCATCTCATCTCATCTCATCTCATCTCATCT 460
236 GGAAGCATCTCATCTCATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 292
461 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 520
293 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 352
521 TGGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 580
353 TGGCAATCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 412
581 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 640
413 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 472
641 TACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 700
473 TACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 532
701 AGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 760
533 GCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 588
761 AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 820
590 AAGCGCGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAGCGGAG 649
821 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 871
650 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 709
872 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 931
710 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 769
932 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 991
770 CTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 829
992 CCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1051
830 CCAAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 889
1052 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1111
890 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 949
1112 CCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1171
950 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 972
1172 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1231
973 ----- 972
1232 GTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1291
973 ----- 972

973 ----- 972
1352 AATTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1411
973 ----- 1026
1412 CCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1471
1027 CCGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
1472 AGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1531
1087 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
1532 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1591
1147 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1206
1592 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1651
1207 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1266
1652 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1711
1267 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1326
1712 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1771
1327 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1386
1772 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1831
1387 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
1832 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1891
1447 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1506
1892 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1951
1507 GGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1566
1952 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1983
1567 TTAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1598

RESULT 2
US-08-560-398-3
Sequence 3, Application US/08560398
Patent No. 597082
GENERAL INFORMATION:
APPLICANT: O'Neill, Sharmen
TITLE OF INVENTION: Nadeau, Jeanette
TITLE OF INVENTION: O'Neill-Specific Gene Expression
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,398

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.rge

Page 14

ILARMILPHPHOSKYVEBLDVAAGKARAVEBDVAMTILPAVKEVIRLHPGPFIL
 WARKSLNDTIDGHYRPAGITAMATVNLGRDPHYWKDLBNPERRPTAGDAEBSI
 LGSPLPLAPFGSGRRACPGKTLGMATVNFVASLHPEWVPSDEKGYDITTEVLKLS
 EWANPLIVKVARPRG"

BASE COUNT	369 a	457 c	407 g	378 t
ORIGIN				

Query Match	40.9%	Score	651.6;	DB	8;	Length	1611;
Best Local Similarity	66.5%	Pred. No.	1.1e-186;				
Matches 1025; Conservative	0;	Mismatches	484;	Indels	33;	Gaps	5;

Qy	6	AGCGAAACCAACCTTGGCTCTTCCCTCCTCGCGCGTCAACATATCTGGCTCGCAATCT	120
Db	74	ACCCAGAAACCACTTGCATGAGGTCCTTTGATCATGCGGCGCACTCGGTAAACATGACT	133
Qy	121	CTCTCTTATGACCTATTCGGGGTGAACCTGTGGGGGGAAAATACCTCTTGGCGCGGTA	180
Db	134	TTTATATACCTGGTCAACCCCGGGTGTCTGCTCGGGGCAAGTACTAACCC-----TA	186
Qy	181	ATATCGGCTTCAACAAACCCGGAACGGTATATCCGGGTCAAAAGGCTTCCCTTGGTT	240
Db	187	CTCTCCCCCTTTGCA-----TCAATTCGGGTCCCAAAGGCTTCCCTCTTATTT	235
Qy	241	GGAAGCATGCACTCATGTCAAGACCTGTAGCTACCGAAGCAATGGCTGATGACGTAG	300
Db	236	GGAAGCATGGGCTCATGACTTCC---CTGAGCCATACCGTATCGAAGCCGCGCGCC	292
Qy	301	AAATTCGAGCCAAAGAGCTCATGGCTTTGACTTGAAGAAGACTCGCGTATGCTACG	360
Db	293	ACATGAGAGCCAAAGGCTCATGGCTTTGACTTCGCGCAACAAGTGTCACTGTACG	352
Qy	361	TGCAATCCGACGTATGCGAAGATTTGAAATGCGCGGTTTTGCTGATGACGAGTT	420
Db	353	TGCAATCCGACGTATGCGAAGATTTGCAACAGCTCGTCTTCCGCGATGTGCCGTC	412
Qy	421	AAAGATCGGCTTACTCACTAGTGTTAACAGGCAATGGTTTTGACACACGGGTT	480
Db	413	AAAGAATCGGCAATACGCTCATGTTTAAACGCGCCATCGGCTTGCCTTACGGAATT	472
Qy	481	TACTGGGAACGCTTGGCGGTATCGCTTCGAAACCATCTTTTGTGTCAAAACAAATGCA	540
Db	473	TACTGGGAAGCTTCAGAGAAATGCGCTCTATACCGCTTTCGCCCCCGCCAGATAAA	532
Qy	541	AGAGCCGAGCCGAAGAGATGATCTCAACCGAATGGTTGATGTTCTTGAAAAACAG	600
Db	533	GCTCTAGGCTCCAAAGCTCTCAATATGCGCGCCAAATGGTCCACATCT---AAATAC	589
Qy	601	AGTATACGAAACCTGTTTGTGTGTGATGTTGCTTAAACCGCGTTCGCTTACAAACATG	660
Db	590	AAAGCCACCGCAGCTTACGTGTTCGCCAAGTGCATAAAAGGCTTCGCTCAGTAAACATG	649
Qy	661	ATGTGCTGTGATTTCCGACAAAGATATAGCTGAAAAAACA-----TGTGAG	711
Db	650	ATGTGCTCGGTGTGACAAAGTATAGCTGCAACCCAAACAGCGGAATGGAAGAC	709
Qy	712	TTACGTGAAATGGTTCGAAGAAGGTTATGATTTGCTCGAAAGTTGAATTGGACTGATAC	771
Db	710	CTTGGAATATATTAGTGAACCAAGGTTATGACGTGTGGGCGGTTAATTGGGCCACAC	769
Qy	772	CTTCTCTGAGCTATCCGAGTTGATCTCTAAAGACGCGGTTGTGAATGTTCAACACTCGTA	831
Db	770	CTTCTCTTTTCTTGCAACATTTCAAGCCCAAATATCGGTTCAAGTGTCTCAACTCTGTC	829
Qy	832	CCAAAGGTAAACCGGTTGTATCCGGATATATCCGAACACGTAATCAACCGGTGAT	891
Db	830	CCCAATGGTGAACCGTTTGTCTCGGCAACAATACGTGTGAACCCAGCTGTATTAACCGAA	889
Qy	892	TTGCTCTGTATTTCCGACAGTTTGTGCTCTCCCTCCATCGGTTCAATATAATTATCCAG	951
Db	890	ACCAATGTGATTTTGTGACGTCTTGTCTCTCTCCCGGAACCTGATCAATTATACAGC	949
Qy	952	CCGGAATTAATGCCCGTCTTTTGGAGATGATATTCAGAGAACAGACACATGTGGGCTC	1011

Db	950	TCGACATGATCGCTGTAATCTTTGGGAAATGATATTCAGAGGACCGGACCGGT	1009
Oy	1012	TTAATCAGATGATCCCTCGCTAGAGTGTCTCATCCAGATATCATCAACGATACAA	1071
Db	1010	TGTATAGATGATATCTCGGAGATATGGCGCTTCATCTCTCATATGTCATATCCAAAGTTCA	1069
Oy	1072	AACGAGCTGAATCAAGTATGTGGGAAATCAAGACCTTAGATGAATCTGACTTGGCTCA	1131
Db	1070	GAGGAGCTAATATGCAATGTGTTCGAAAAAGCAGCGCGCTGCGAAGAAATGAAGTGGCAGTG	1129
Oy	1132	CTTCCATATCTAATCGGCTGTGTGTAAGAAGTATTAAGGCTTCATCTCTCAGGCCACTT	1181
Db	1130	ATATAGTATCTATCCAGACGGTGTGTATAGAGATGTCTGGCTGCATCCGCGCGGCCACTT	1189
Oy	1192	CTATCATGGGCGCGTTTGGCCATTAACAGACAGATGTGTGATGTGTCTGTTCGGCA	1251
Db	1190	CTATCATGGGCGCGGTGTCTCATATGAATCAACACATGAATGGGTATCAAGTATCTGGC	1249
Oy	1252	GGAACACAGCATATGTTAACTATGGGCGGTATCCATATATCCAACAGTGTGGGTAT	1311
Db	1250	GGAACACATGATATGTCAACAGTGGGCTATTGACAGGAGCCACAGTGTGGAAGAC	1309
Oy	1312	CGTTTGGATTTAAACCTGAAGGTTGTGTGGCAAAAGAAAGTGAAGTATTTCCGTT	1371
Db	1310	CCAATCTGAATTTATGCCGAGAGGTTGTCACTCGCGGTGGAAGATGCCAATTTTCGATA	1369
Oy	1372	CTTGGGTCCGATTTTGAAGCTTGCACTTTCGGGTCCGGTCCGTCCGATTTGGCCCCGGAG	1431
Db	1370	CTCGGGTCCGATTCGAAGCTTGCTCCATTTGGGTCCGGTATGAGAACCGTCCCAAGGAG	1429
Oy	1432	AATCTTGTATTACTACCGCTATATGTTTGGACGGCATATGATCTAATATGAATTTGAATGG	1491
Db	1430	ACTCTTGAATGGGCTACAGCTAATCTTTGGGTGGCGTCTCTTGATATAGTTCCAAATGG	1489
Oy	1492	GGAACGTCCGATGTATACGGGCTTGACTTATCTGAGAACTGAAGCTTTCTTGCAGATATG	1551
Db	1490	GTAACGCTGATGAGAAAGGCTGTGATCTGACGGAAGGTGAAGTCTCTAGTGAATG	1549
Oy	1552	GCTAATCTCTTCCGTGAATATGGCGGTATAGGCGCAAGTAA	1593
Db	1550	GCTAACCTTCTACCGCTGAATGTGCCCCGAGCGGTGATATA	1591

RESULT 9	740 bp	DNA	linear	PLN 29-MAR-2002
ATH552847/c				
LOCUS				
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, left border, clone			
ACCESSION	AF552847			
VERSION	AJ552847			
KEYWORDS	AJ552847.1 GI:29368998			
SOURCE	left border; T-DNA flanking sequence.			
ORGANISM	Arabidopsis thaliana (thale cress)			
	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;			
	rosids; eustosida II; Brassicales; Brassicaceae; Arabidopsids.			
REFERENCE	1			
AUTHORS	Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., Dehose, R., Pelletier, G., Lepiniec, L., Caboche, M. and Lecharoy, A.			
TITLE	T-DNA integration into the Arabidopsis genome depends on sequences of pre-infection sites			
JOURNAL	EMBO Rep. 3 (12), 1152-1157 (2002)			
MEDLINE	12363535			
PUBLISHED	22446565			
REFERENCE	2 (bases 1 to 740)			
AUTHORS	Balzergue, S.			
TITLE	Direct Submersion			
JOURNAL	Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue			
COMMENT	Gaston Crenieux, 91057 Evry cedex, FRANCE			
	PCR was performed on DNA from transfectants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from			

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 00:02:34 ; Search time 5880 Seconds

(without alignments)
11083.169 Million cell updates/sec

Title: US-10-022-025A-2

Perfect score: 1593
Sequence: 1 atggctacgaactcgaag.....tgcgcgcagcgcagcttaa 1593

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 2888711 segs, 20454813386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Genemb1:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_cm:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pin:*
35: em_htg_rtd:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_hnho_hum:*
40: em_hnho_mus:*
41: em_hnho_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	1360	85.4	55870	8 AC005819	AC005819 Arabidops
C 2	1360	85.4	54825	8 AC006418	AC006418 Arabidops
C 3	976	61.3	1605	8 BT008891	BT008891 Arabidops
C 4	976	61.3	1824	8 AY062748	AY062748 Arabidops
C 5	976	61.3	1902	8 AB036059	AB036059 Arabidops
C 6	975.6	61.2	1878	8 AY086928	AY086928 Arabidops
C 7	794	49.8	142001	8 ATE21F14	ATL38642 Arabidops
C 8	651.6	40.9	1611	8 AF022463	AF022463 Glycine m
C 9	575.8	36.1	740	8 ATH552847	ATH552847 Arabidops
C 10	474.4	29.8	86436	8 AC007323	AC007323 Genomic s
C 11	474.4	29.8	104001	8 AC023628	AC023628 Arabidops
C 12	355	22.3	1799	6 AR070567	AR070567 Sequence
C 13	355	22.3	1799	6 PSU34744	PSU34744 Phalaenopsi
C 14	348.2	21.9	1990	8 AF049067	AF049067 Pinus rad
C 15	336.2	21.1	565	11 AL772960	AL772960 Arabidops
C 16	295	18.5	66770	8 AP004704	AP004704 Oryza sat
C 17	289	18.1	2158	8 BT000225	BT000225 Arabidops
C 18	289	18.1	2158	8 AY136401	AY136401 Arabidops
C 19	260.4	16.3	504	8 ATH553770	ATH553770 Arabidops
C 20	239	15.0	108387	8 AC016662	AC016662 Arabidops
C 21	237.4	14.9	2087	8 MZCP450A	L33209 Zea mays cy
C 22	236.4	14.8	310	11 AL772952	AL772952 Arabidops
C 23	228.4	14.3	114477	2 AP005682	AP005682 Oryza sat
C 24	218.2	13.7	1668	8 AB096259	AB096259 Oryza sat
C 25	205.6	12.9	122218	8 AC099399	AC099399 Oryza sat
C 26	205.6	12.9	138831	8 AC140005	AC140005 Oryza sat
C 27	198.4	12.5	79054	8 P21F23	AC027656 Sequence
C 28	194.6	12.2	471	11 AL772558	AL772558 Arabidops
C 29	174.6	11.0	82390	8 AB016893	AB016893 Arabidops
C 30	173.2	10.9	153247	8 AC137992	AC137992 Oryza sat
C 31	155.6	9.8	124573	8 AP005175	AP005175 Oryza sat
C 32	144.2	9.1	1704	8 AF022458	AF022458 Glycine m
C 33	143.6	9.0	169270	8 AC083943	AC083943 Genomic s
C 34	143.6	9.0	324367	8 AE017091	AE017091 Oryza sat
C 35	141.8	8.9	1789	6 AX028817	AX028817 Sequence
C 36	141.8	8.9	1789	6 AR155332	AR155332 Petunia x
C 37	135.6	8.5	1667	6 AX028838	AX028838 Sequence
C 38	135.6	8.5	1764	6 AB057673	AB057673 Torenia h
C 39	134.6	8.4	1545	8 AY082612	AY082612 Ocimum ba
C 40	134.4	8.4	39601	8 AC079678	AC079678 Arabidops
C 41	133	8.3	1554	8 AY082611	AY082611 Ocimum ba
C 42	131.8	8.3	1542	8 AF271649	AF271649 Arabidops
C 43	131.8	8.3	1542	8 AF271650	AF271650 Arabidops
C 44	131.8	8.3	1542	8 AF271651	AF271651 Arabidops
C 45	130.2	8.2	1748	8 AF313491	AF313491 Matchiola

ALIGNMENTS

RESULT 1
AC005819/c
LOCUS
DEFINITION
AC005819
Arabidopsis thaliana chromosome 2 clone T3A4 map C1C6C03, complete
sequence.
AC005819
VERSION
AC005819.3 GI:20197448
KEYWORDS
HTG.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidops.
REFERENCE
1 (bases 1 to 55870)
Rounsley,S.D., Lin,X., Kaul,S., Shea,T.P., Fujii,C.Y., Mason,T.M.,

JOURNAL
 2
 Unpublished
 Sheng, M., Romung, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.
 2 (bases 1 to 55870)
 Lin, X.
 Direct Submission
 Submitted (09-Mar-2000) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 3 (bases 1 to 55870)
 Town, C.D. and Kaul, S.
 Direct Submission
 Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, cdrom@tigr.org
 On Apr 18, 2002 this sequence version replaced by: 6598489.
 COMMENT
 FEATURES
 source
 1..55870
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="2"
 /map="C106C03"
 /clone="T3A4"
 1..14747
 /note="overlap with BAC clone F13A10
 AC006418:70079..84825."
 92..1276
 /gene="At2g46630"
 /note="synonym: T3A4.1"
 <92..>1276
 /gene="At2g46630"
 92..1276
 /gene="At2g46630"
 /codon_start=1
 /product="putative extensin"
 /protein_id="AAC69930.1"
 /db_xref="GI:3831447"
 /translation="MSNRPRVRLSSTAPRTAGSSDPPEPPQPPPEPTSSSLVSPAPK
 QSSPRQQRSSPPRODPPSPPOQCPPTPRQQAPTSPQSSPSPPSSKHS
 PTPPKAPPPPPPPRSITSPSPKVSQALPPKSPSPSPASSTSSSEKTR
 SSSSENRKAPSPRVLSPTSLPASHLSRETQKLLTAKTKQTETRNHNQNV
 HDYNNYNNQNSYNQNNQNNPKRRQSSSDSENISTVITTAENNGAVN
 ELRPPQNKTKGSGSTHSRVSHGTGKRLQSSSSSDSGEKKTKTKVPRKN
 NLPKAFMNSNVQNNINNSIYVNSTASHDPGHLKISRKPSGSDMGFVSDYGNNGY
 TN"
 723..810
 /rpt_family="CACCATn"
 1345..1382
 /rpt_family="AT-rich"
 complement(1904..1925)
 /rpt_family="AT-rich"
 3740..5558
 /gene="At2g46640"
 /note="synonym: T3A4.2; predicted by genSCAN and
 GeneIndex"
 join(1<3740..4072,4525..4839,5083..5215,5407..5558)
 /gene="At2g46640"
 join(3740..4072,4525..4839,5083..5215,5407..5558)
 /gene="At2g46640"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAC69921.1"
 /db_xref="GI:3831438"
 /translation="MYKEDIVGRKNKKKNEGISEIKNTKAILDQVGLVDALDNME
 DQVLITGFGEPTLNFKEBEDDDDECGSDLYVVIDSIIKNVQSSDPIISVNNK
 PDHEDVADTDHSGSIKTAEPTAVTALEAIEPQKRTILAEIFMEDHDKDQV
 HCKKNRNPLDGEVRYKQNGSKLSPFVFKKLMKSKSEENDLRFIKKRNQMI
 KMLKKKHIPVDATKLRKQDVPYKPTRNCEALSHSYHLKIQEGCIWNTVMCFLL
 NITKLTVFVPMHIVFYVYICFLITYINAP"
 6216..7156
 /gene="At2g46660"
 /note="synonym: T3A4.3; supported by full length cDNA:
 GenBank:3743"
 join(3716..6447,6552..6618,6704..7156)
 mRNA

CDS	/gene="At2g46650" join(6370..6447,6552..6618,6704..6957) /gene="At2g46650" /codon_start=1 /product="putative cytochrome b5" /protein_id="AAC69923.1" /db_xref="GI:3831439"
repeat_region	/translational="ANLISFHDVAKRKCNDCMILHGRVYDISTFMDEHPCGDNVL LATGDAADIDEDVNHSKDKELMKRYCIGDVDOSTVFVQOYIPEWEKSTAAETT KEESGGKLIVIPLLTGAVAFARFPNNK" complement(7227..7247)
repeat_region	/rpt_family="AT_rich" 7693..7715
repeat_region	/rpt_family="AT_rich" complement(7799..7822) /rpt_family="AT_rich"
gene	complement(8165..9980) /gene="At2g46660" /note="synonym: T3A4.4" complement(join(<8165..8782,9006..>9980)) /gene="At2g46660" complement(join(8165..8782,9006..9980)) /gene="At2g46660" /codon_start=1 /product="putative cytochrome P450" /protein_id="AAC69923.1" /db_xref="GI:3831440"
mRNA	/translational="WALIKBSLSLIFALLSKCSITSGTNLAIBSLAVTIWLAISEPM TYGVGMANKYIFGRLLSGLTKGTNVPGKGPLVGSMISLSTLAHRADAEKF GARLRALFLAGTRVTCNDPAKEILNSPVFAPIVEKSSLMFRRLAPGAPEV WMYTLRIASNLTFSTQIRRAFORRVVISQVPERLEKQSNECFVRBLTKYASLN NMACSFGCEIEKNHTELEWEEBEGVDLGLNTDHLFWLSDFORARSCTPL VPKXNPFYRIISEHNQGTDLPRDPFDYLISHGSDKLDPDIIAYLVMEIPEGDTL VALILMARLVLPDMOSTVQNLEDQVVGKSRLDESDDLASYLAAYVEYLRLH PVALISMARLALITDIWDGLVPAQTVMAMVAASHDPHYVDDPLEFKERVAKE GESEVSFGSDLRALPESGGRICPGNKLGFTTWMTMMHEEMQSPDGNGVDLS EKRLSCPMANLPLKLRRRRS" 10124..10180
repeat_region	/rpt_family="AT_rich" complement(14979..15002) /rpt_family="A)n"
repeat_region	15546..15568 /rpt_family="AT_rich" 17214..17366 /rpt_family="AT_rich"
repeat_region	17478..17503 /rpt_family="(A)n" complement(17523..17543) /rpt_family="AT_rich"
repeat_region	complement(17647..17688) /rpt_family="AT_rich"
repeat_region	complement(17874..17894) /rpt_family="AT_rich"
repeat_region	complement(18001..18060) /rpt_family="AT_rich"
repeat_region	complement(19152..19796) /gene="At2g46670" /note="synonym: T3A.5; predicted by geneScan and genefinder" complement(join(<19152..19310,19344..>19796)) /gene="At2g46670"
mRNA	complement(join(19152..19310,19344..19796)) /gene="At2g46670" /codon_start=1 /product="hypothetical protein" /protein_id="AAC69924.1" /db_xref="GI:3831441"
CDS	/translational="MYLCRFEESKAEXAVNALREESTSGEPKPTSEHKLRKYARD CGSSTIGASSSHRKEREESKKQMSQSBALKMPKLRKDRCFTDKHLKIQDOT DSWEIKRNQLQRFVOSRKKLAEQRFRVKGQFRTVNSASTKS" 19989..21343 /note="molecular marker Atbh7"

Gene	complement (20037..21512) /gene="At2g46680" /note="synonym: T3A4.6; identical to SP:P46897; supported by cDNA: gi_15027938_gd_AY045826.1" complement [join(20037..20039,20111..20756,20977..>21512)] /gene="At2g46680" complement [join(20340..20756,20977..21336)] /gene="At2g46680"
CDS	/codon_start=1 /product="homeodomain transcription factor (ATHB-7)" /protein_id="AAC69925.1" /db_xref="GI:3831442" /translation="MTGGGEYSPAMWSABPFLTMKQKKKKNNKNNORRFSDEQIKSL EMPFSSRIERPRKKQLARIEGICAPQVATWKNRRARRKSKOLETEYNILROYND LASQESLAKKKQKLVSBLORLEATQKQGEERQSCDQAVNLSSHHSEEN RRRKEEYRPPMEMKDKGHHGVACDHDVEDDNGISNNIKETFTGGFEEPPDLAM IVEPADSLTSSDDMRGKSDTTLIDQSSNNYPMDPFWSS" complement (26336..22664) /rpt_family="AT_rich"
repeat_region	23532..23571 /rpt_family="AT_rich"
repeat_region	23760..23783 /rpt_family="AT_rich"
repeat_region	24595..24656 /rpt_family="CAAAA)n" complement (25135..25155) /rpt_family="AT_rich"
repeat_region	25903..25923 /rpt_family="AT_rich"
repeat_region	complement (26667..26771) /rpt_family="AT_rich"
repeat_region	complement (26830..26901) /rpt_family="R:ATREP1 ATREP1 ATREP1 dispersed repeat from GB D86608 - a consensus." 30015..30038 /rpt_family="(TAAAA)n"
repeat_region	complement (31911..31933) /rpt_family="AT_rich"
repeat_region	complement (32444..32466) /rpt_family="AT_rich"
Gene	35278..35832

Query Match	85.4%	Score 1360;	DB 8;	Length 55870;
Best Local Similarity	87.7%	Pred. No. 0;		
Matches 1593;	Conservative	0;	Mismatches	0;
			Indels	223;
			Gaps	1;

QY	I	ATGGCTACGAAACCTCGAAACCTCCTTAAATCTCTTGACCTTTGTGTCGAAATGACAGGTTCTTA	60
Db	9980	ATGGCTTACGAAACCTCGAAACCTCCTTAAATCTCTTGACCTTTGTGTCGAAATGACAGGTTCTTA	9921
QY	61	AGCGAAACGCAACCTTGACTTCTCCCTCCTCGCGCTGCACAATTCATCTTGCTCGCATATCT	120
Db	9920	AGCGAAACGCAACCTTGACTTCTCCCTCCTCGCGCTGCACAATTCATCTTGCTCGCATATCT	9861
QY	121	CTCTTCTTATGACCTTATCCCGGTGACCTTGCTTGAGGGAATACTCTTTCGCGCGGTTA	180
Db	9860	CTCTTCTTATGACCTTATCCCGGTGACCTTGCTTGAGGGAATACTCTTTCGCGCGGTTA	9801
QY	181	ATATCCGGTTCATACAAACCCGAAAGATTATTCGCCGATCCAAAGGCTTCCCTTTGGTT	240
Db	9800	ATATCCGGTTCATACAAACCCGAAAGATTATTCGCCGATCCAAAGGCTTCCCTTTGGTT	9741
QY	241	GGAAGCATATGACATCATGTCAAGACCTTAGCTACCGACGAAATCGCTGATGACAGTGAG	300
Db	9740	GGAAGCATATGACATCATGTCTAAGCATCTTAGCTACCGACGAAATCGCTGATGACAGTGAG	9681
QY	301	AAATTCGGAGCCGAAGAGGCTCATGGCTTTCAGCTTAGAGAGAGACTCGCGTATGTCACG	360
Db	9680	AAATTCGGAGCCGAAGAGGCTCATGGCTTTCAGCTTAGAGAGAGACTCGCGTATGTCACG	9621
QY	361	TGCATATCCGACGTAGCGAAGAGATTCTGAATAGCCCGGTTTTTGTCTGATCGACCGGTT	420
Db	9620	TGCATATCCGACGTAGCGAAGAGATTCTGAATAGCCCGGTTTTTGTCTGATCGACCGGTT	9561

QY	421	AAAGAAATCGGGTACTCTGATCTGATGTTTAAACAAGCAATTTGGTTTGGACCAACGGTGT	480
Db	9560	AAAGAAATCGGGTACTCTGATCTGATGTTTAAACAAGCAATTTGGTTTGGACCAACGGTGT	9501
QY	481	TACTGGCGAAGCGGTCTGGCGGTATCGCTTGAAACCATCTCTTTAGTACAAACAAATTCGA	540
Db	9500	TACTGGCGAAGCGGTCTGGCGGTATCGCTTGAAACCATCTCTTTAGTACAAACAAATTCGA	9441
QY	541	AGAGCCGAGACCGCAACGACGATGATCTCAAGCCAGATGGTTGAGTTTGTGAAAAACAG	600
Db	9440	AGAGCCGAGACCGCAACGACGATGATCTCAAGCCAGATGGTTGAGTTTGTGAAAAACAG	9381
QY	601	AGTGTACAGAACCTGTTTGTTCGTGATGTTGCTTTAAACGGCGCTGCTTAACACATG	660
Db	9380	AGTGTACAGAACCTGTTTGTTCGTGATGTTGCTTTAAACGGCGCTGCTTAACACATG	9321
QY	661	ATGTGCTCTGATTTCCGACAGAGATGATGAGCTTGAAAAAACATGTGATGATGATGAA	720
Db	9320	ATGTGCTCTGATTTCCGACAGAGATGATGAGCTTGAAAAAACATGTGATGATGATGAA	9261
QY	721	ATGTGTCGAGAGGTTATGATTTGTTCGGAACGTTGATTTGACATGATCACTTCTTGG	780
Db	9260	ATGTGTCGAGAGGTTATGATTTGTTCGGAACGTTGATTTGACATGATCACTTCTTGG	9201
QY	781	CTATCGGAGTTGATCTCTCAAAAGCTCCGGTCTGATGATTTCCACACTCCGACCAAGATA	840
Db	9200	CTATCGGAGTTGATCTCTCAAAAGCTCCGGTCTGATGATTTCCACACTCCGACCAAGATA	9141
QY	841	AACCGGTTGTATCCCGGATTTATATCCGAAACCCGTAATCAACCGGTGATTTGGCTCGT	900
Db	9140	AACCGGTTGTATCCCGGATTTATATCCGAAACCCGTAATCAACCGGTGATTTGGCTCGT	9081
QY	901	GATTTGCGTCGACGTTTGTGCTCTCCCTCAGATGTTACAGTAAATATTCGACCCGACATA	960
Db	9080	GATTTGCGTCGACGTTTGTGCTCTCCCTCAGATGTTACAGTAAATATTCGACCCGACATA	9021
QY	961	ATCGCGGTCTCTTG-----	974
Db	9020	ATCGCGGTCTCTTG-----	8961
QY	975	-----	974
Db	8960	TGAAAAATGCTTAGGATTTATTAAGATTAAAAATGAAATTTGAGACTCAATGTGACGTGT	8901
QY	975	-----	974
Db	8900	TGTGGAAATTAAGACATTAAGATTTTGTGACATCAATTAAGTAAATTTAGAT	8841
QY	975	-----GGA 977	974
Db	8840	TTTATTAACAGTTTCTATTAATAAAGTAAAAAGTAAATTTTGAAGTATTAATTAACGGA	8781
QY	978	GATGATATTACAGAGAACAGACAGCTGGCGGTCTTAATCGATGATCTCTCGTACGAT	1037
Db	8780	GATGATATTACAGAGAACAGACAGCTGGCGGTCTTAATCGATGATCTCTCGTACGAT	8721
QY	1038	GGTCCCTCATCAAGATATGCAATCAACCGTCAAAACAGAGCTGATCAAGTATCGGAA	1097
Db	8720	GGTCCCTCATCAAGATATGCAATCAACCGTCAAAACAGAGCTGATCAAGTATCGGAA	8661
QY	1098	ATCAAGAGCCCTAGATGAATGATGACTTGGCTTCACTTCATATCTAACCGGTGTGGTGA	1157
Db	8660	ATCAAGAGCCCTAGATGAATGATGACTTGGCTTCACTTCATATCTAACCGGTGTGGTGA	8601
QY	1158	AGAAAGATTAAGAGCTTCACTCCAGAGCCCACTTCTATCATGTGGCCGCTTTGGCCATAC	1217
Db	8600	AGAAAGATTAAGAGCTTCACTCCAGAGCCCACTTCTATCATGTGGCCGCTTTGGCCATAC	8541
QY	1218	AGACACGATGTTGATGTCGTCTTGTTCGGCAGAGACCAACAGAAATGGTGAACATGTG	1277
Db	8540	AGACACGATGTTGATGTCGTCTTGTTCGGCAGAGACCAACAGAAATGGTGAACATGTG	8481

QY	1278	GGCGGATCGCATGATCCACACCGTGGAGTTGATCCCTTGAGAGTTAACTGAGAGTT	1337
Db	8480	GGCGGATCGCATGATCCACACCGTGGAGTTGATCCCTTGAGAGTTAACTGAGAGTT	8421
QY	1338	CGTGGCAAAAGAGAGTGAAGTGAAGTTCGGTTCCTTGAGTGGATTTGAGACTTGAC	1397
Db	8420	CGTGGCAAAAGAGAGTGAAGTGAAGTTCGGTTCCTTGAGTGGATTTGAGACTTGAC	8361
QY	1398	TTTCGGGTCGGGTCGTGGATTTGTCGCCCGGGAAGAACTTGGTTTACACCTTATGTT	1457
Db	8360	TTTCGGGTCGGGTCGTGGATTTGTCGCCCGGGAAGAACTTGGTTTACACCTTATGTT	8301
QY	1458	TTGACGCGCATGATGTTACATGATGATTGAATGAGGACGTCGCCATGATTAAGCGCGTTGA	1517
Db	8300	TTGACGCGCATGATGTTACATGATGATTGAATGAGGACGTCGCCATGATTAAGCGCGTTGA	8241
QY	1518	CTTATCTGAGAACTGAGGCTTTCTTGGAATGGCTATCTCTTCCTGTAATTCG	1577
Db	8240	CTTATCTGAGAACTGAGGCTTTCTTGGAATGGCTATCTCTTCCTGTAATTCG	8181
QY	1578	CCGTAGCGCGCATTA 1593	
Db	8180	CCGTAGCGCGCATTA 8165	
RESULT 2			
AC006418/c			
LOCUS	AC006418	84825 bp	DNA linear
DEFINITION	Arabidopsis thaliana chromosome 2 clone F13A10 map C10C06C3,		
ACCESSION	AC006418		
VERSION	AC006418.4		
KEYWORDS	GI:20197771		
SOURCE	HTG.		
ORGANISM	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabidopsis thaliana (thale cress)		
TITLE	Arabidopsis thaliana (thale cress)		
REFERENCE	Arabidopsis thaliana (thale cress)		
AUTHORS	Arabidopsis thaliana (thale cress)		
JOURNAL	Arabid		

RNA
 complement (join(<1047. .1322,1409. .1518,1595. .1597. .2074. .2431,2503. .2828,2916. .3029))
 /gene="At2g46480"
 CDS
 complement (join(1047. .1322,1409. .1518,1595. .1597, 2074. .2431,2503. .2828,2916. .3029))
 /gene="At2g46480"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAD20159.1"
 /db_xref="GI:4415923"
 /translation="MTDACCILKQNEKQKVRPFQHTVIGAFNDTEMLHSLRQEK
 RLEENALMDDSDLOKLETAAMARSRSVDSAPLGNVTIWNQERYRGSFEDMLQAD
 QIIMARYVSGAKTNKTLAHQELIETQMLAAEESTDIDQCRVLDLRDGLQIA
 RAHQIPECKLITNKLRLAMQIWEDELENGQITFTLOASRALPAILICLIRLNL
 EYHLILPKNRPFPRRRLINPRLKYLVALESDNLAAVYVNSIVNANQSDSRVFLH
 TDKLAQMSMFLINPGERATTHVQCFPEPTNLSGVSFVLSAAMKKPYFKTA
 RSEVSGSENLKTRIFRINSMNLNHLFTEPRFLPEKLFLPDDVIVVQILGLTP
 IDLGVKSNEDPFRFCGMAVGMNIFDLKERRKNITETITFQNTLENRTLMGLTP
 PGLTFVYLTQPLRKMLLGLGDKGIDVKKIERSAVIHNGHMKPTEWGISKTP
 YTKYTNDDHPIYFTCRLEF"
 4060. .4696
 /gene="At2g46490"
 /note="synonym: F13A10.2"
 /db_xref="GI:4415923"
 /translation="MAADGIFRSIFEGCISGLDLSIRRRPHKNGGALHDKSGAGK
 NQNRQPRSRHSGSSSEIFPIRKSSEGNITAMNLFSSSSSSNNQSLSSSSSLNN
 LASDLPVNAATEEPESSRNKQLRWTIGEDD"
 complement (4483. .4517)
 /rpt_family="AT-rich"
 complement (6756. .6792)
 /rpt_family="AT-rich"
 complement (11207. .13602)
 /gene="At2g46500"
 /note="synonym: F13A10.3; similarity to ubiquitin family
 of proteins; supported by cDNA: g1_16930424_gb_Af419566
 1_Af419566"
 complement (join(<11207. .11764,11858. .>13602))
 /gene="At2g46500"
 complement (join(11384. .11764,11858. .13177))
 /gene="At2g46500"
 /codon_start=1
 /product="expressed protein"
 /protein_id="AAD20161.1"
 /db_xref="GI:4415931"
 /translation="MSSAGVALSPVRSBPLIMPLVRANSCIDSYEDDTITLYLIPGS
 VTPMRLESDSESVKLRIQSYRGFVVNRQKLVPQGRLASNSNRDYGVSSEGLIL
 LVTKSLDVLVLTQTCGKCFHYERGNIGVYKQISKRGDVPDDEBELLYEG
 KLEQSLINDICRNDSDVLHLVRSAAKRVFVPEKNEELSIAPQADKKGRKXS
 VPKKLSLPPVYNBKAQVPLVYKMIQSDGLKSGNSPVASSEGTGAAFMQGSPS
 NKVGFATKIDEEPAENNPQGLPLSPNGEIKKQTKYGEGLAREVAAYILDHPSG
 KSGEISIEPAGVPPTAMTECLHPGNPKIKTKITGLIQWTEPDGSCDMGPEL
 VSEVHGIIVLDIRLANADHGNILMTKQESKLVLVPIIDREYCIPISEFECTPEWL
 WEPARPYAETQETVIRSLDAEDIDLTFHGMKVPASTAQTIRLSTYLLKGYRG
 TAPETILCRETILSKSLVEEMVEBAQEAIVLPTSEAAFLDALSDVMDIHLDSQET
 "

```

/db_xref="GI:20197780"
/translation="MSLRRTIIDLRLRQTRATLSKSLPFSMSHLSASAGATAVRNPLG
RDLSSIPFAOAKLKPDSNTLVTDRSISSISQSEELNRAKFSRQSSRGYTNCSFLR
KIPVPHIHIEGKEITLADYVHOEMRNILVMSLGFQIIVLKDITLFL"
gene
    complement(15830..17530)
    /gene="At2g46510"
    /note="synonym: F13A10.4"
    complement(<15830..>17530)
    /gene="At2g46510"
    complement(15830..17530)
    /gene="At2g46510"
    /codon_start=1
    /product="putative bHLH transcription factor"
    /protein_id="AA020162.2"
    /db_xref="GI:20197775"
/translation="NMMSDLGWDDEDEKSVVSAVLGHLASDPLRANSNQNLEFLVMT
DPTLNKSLSLVDMPNSENFSWNRYAIFMOQTMSRSGQVLDGDCCEPNEEESKY
VRSYNNPNMGAREETWQDRKRVLOKHLFGSDSDNNALSEKVTATEIFFLAMT
FPRNHEGGRGRCYSKSKVNLSDANSRSDYCFRSPAKSGIRITIVVPTDAPVLE
LGSVMSLPENIGLVSVQALPFRRTVQPVMTSNTMGTGTHKLGODLSGHAYPK
LEVRNLDERFTPOQSWEGYNNKGPFTFGTQPDVYKVLNNVNNVNNNTKTQLEFA
GSEVAASNPSTVTOQESKESCTEKRPVSLAGAGVSVDEKRPKRGKRNAREE
PLNHEAEORREKLNRFPALRSVVPNSIKMDKASLGDASIKELQEKYINDEB
RVGTDSLSNSNTIVESPEVDIOAMNEEVVVRISPDSPASRIIOAMNSVSL
MEKSLADDTMHTFVIKSNNSDPLTEKXLIIAFPEPSTSTPPLPSSQVSGDI"
repeat_region
    17920..17946
    /rpt_family="GA)n"
    complement(18392..18454)
    /rpt_family="AT_rich"
gene
    21510..24428
    /gene="At2g46520"
    /note="synonym: F13A10.5"
    <21510..>24428
    /gene="At2g46520"
    21510..24428
    /gene="At2g46520"
    /codon_start=1
    /product="putative cellular apoptosis susceptibility
    protein"
    /protein_id="AA020163.1"
    /db_xref="GI:4415933"
/translation="MEWNRRETLVFLSQCLNTLSPIPERPTEALRSLDAPDQNYGL
AVRLVLEPAIDQETRAAANFKNHLRSRMRHAGDSGISPIYDSKEKQIKTLVSLM
LSASPRQSOSEALTVIGADPEKAPVLPPELLNLQVALQKATVAVNGSLVSGAS
SIRKKESEYRTDALFVLDKYCLDNPAALTEIFLTKSLSIDSAASGSPPLIKFL
ESQRCTCTIYSLNFODLPPEPDHMKEMWGEPRKYLASNVPALSTEEELTVLDLR
AAICENINHYIEKNEEFQGLNEFASVVTTLRDVSKSRRQQLATIAIKPTSTST
SVHNLPAAGNVKIEICQSTIVIPNLSLVEDEBI FENYTFEIRDMESKDDTTRRI
ACELKGLATNYKQVTEVVSLEIQRLSSFSANPSANMKDQKCALVIVLSLTKKAG
GASVSTDLIVQNFANIILPELQSRDVSFPMKGSGLTLPFESHKPKPMPLF
PELVRFILASNNVSYAASCIKELILVEKGAGRNVAAGDLSPLLQMTMLPFAL
KPESEENOVLMKIMEVLGADISAEVAPCIQGLISLISECKAPKNIFENHYLE
SVATLVRRACERDILSISAPETSLPSELQMTLNDIIEELPYGCOLLAOLVELNRPTL
SPNTWQIFLLLSPEKSKSGNVPALRLQAPLQKRPHTVQENRSLQVIGIEKLV
ASPTDQGEYFIINTLILENDIVSYAVYMGVNSALTRQNKCTVAFQSLVFNPSL
FLVHGQAYLVETNNTVQPNITAIIVHFVI PMLKILMGSEVXLAIVAAVTRILCELP
ALDPSAKLWGMKLDLIVLVSREPQERVLDEPEMEISENGYTAAPVGLNAAKX
EDPLKDIKDQKFLVASVSRLSASGPRYFQIIGENLQANGTALIQLCNAYNCGIA"
repeat_region
    25088..25183
    /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
    microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
    74A"
repeat_region
    25090..25183
    /rpt_family="GA)n"
    25090..25183
    /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
    microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
    74A"
repeat_region
    25090..25162
    /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
    microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
    74A"
repeat_region
    25090..25162
    /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
    microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
    74A"

```

```

repeat_region
    74A"
    25090..25183
    /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
    microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
    74A"
repeat_region
    25090..25154
    /rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.chaliana
    microsatellite repeat region (clone 74A)#Rf:ATR0046|X93607
    74A"
Query Match      85.4%; Score 1360; DB 8; Length 84825;
Best Local Similarity 87.7%; Pred. No. 0;
Matches 1593; Conservative 0; Mismatches 0; Indels 223; Gaps 1;
QY      1  ATGGCTAGAAACCTCGAAGCTCCTTAATCTTTGACCTTTTGTGCAAAATGACAGGCTCTA 60
DB      80058  ATGGCTAGAAACCTCGAAGCTCCTTAATCTTTGACCTTTTGTGCAAAATGACAGGCTCTA 79939
QY      61  AGCCAAACCAACCTTGCCTTCTCCCTCCTGCGCTCAAAATCATCTGGCTGCCAATCT 120
DB      79998  AGCCAAACCAACCTTGCCTTCTCCCTCCTGCGCTCAAAATCATCTGGCTGCCAATCT 79939
QY      121  CTCTTTATAGACCTATCCCGGTGACCTGCTGGGGGAAATCCCTTCGCGCGGTTA 180
DB      79938  CTCTTTATAGACCTATCCCGGTGACCTGCTGGGGGAAATCCCTTCGCGCGGTTA 79879
QY      181  AATCCGGTTCATCAAAACCGAAGCTTATCCCGGTCCAAAAGGCTTCCCTTTGGTT 240
DB      79878  AATCCGGTTCATCAAAACCGAAGCTTATCCCGGTCCAAAAGGCTTCCCTTTGGTT 79819
QY      241  GGAAAGATGTCATCTCATCTCAAGCACTCTAGCTCAACGAGAAATCGCTGAGCGTGA 300
DB      79818  GGAAAGATGTCATCTCATCTCAAGCACTCTAGCTCAACGAGAAATCGCTGAGCGTGA 79759
QY      301  AAATTCGAGCGAAGAGGCTCATGCTTTCAGCTTGAAGAGACTCGGATGTCACG 360
DB      79758  AAATTCGAGCGAAGAGGCTCATGCTTTCAGCTTGAAGAGACTCGGATGTCACG 79699
QY      361  TGCAATCCCGACGTACGAAAGATTTGTAATAGCCCGGTTTTTGTCTGATCGACGGTT 420
DB      79698  TGCAATCCCGACGTACGAAAGATTTGTAATAGCCCGGTTTTTGTCTGATCGACGGTT 79639
QY      421  AAAGAATGGGTTATCTCATGATGTTTAAAGAGCAATTTGTTTGCACACAGGTTT 480
DB      79638  AAAGAATGGGTTATCTCATGATGTTTAAAGAGCAATTTGTTTGCACACAGGTTT 79579
QY      481  TACTGGCGAAGCTTGGCGGTATCGCTTGAACCATCTCTTAACTCAAAACCAATCGA 540
DB      79578  TACTGGCGAAGCTTGGCGGTATCGCTTGAACCATCTCTTAACTCAAAACCAATCGA 79519
QY      541  AAGCGCGAAGCGCAACGAGCTGATCTCAAGCGAGATGTTGATTTCTTGAAGAAACAG 600
DB      79518  AAGCGCGAAGCGCAACGAGCTGATCTCAAGCGAGATGTTGATTTCTTGAAGAAACAG 79459
QY      601  ACTAGTAAGCAACCTGTTTGTTCGTAGTGTCTTAAACGGCGCTGCTTAAACAATG 660
DB      79458  ACTAGTAAGCAACCTGTTTGTTCGTAGTGTCTTAAACGGCGCTGCTTAAACAATG 79399
QY      661  ATGTGCTGTGATTTGAGACAGAGTATGAGCTTGAAGAAACCAATGTTGATGAGTAA 720
DB      79398  ATGTGCTGTGATTTGAGACAGAGTATGAGCTTGAAGAAACCAATGTTGATGAGTAA 79339
QY      721  ATGTGGAAGAGGTTATGATTTGCTCGAAGCTTGAATTTGATCTGATCACTTCTTGG 780
DB      79338  ATGTGGAAGAGGTTATGATTTGCTCGAAGCTTGAATTTGATCTGATCACTTCTTGG 79279
QY      781  CTATCGAAGTTGATCTCAAGACTCGGCTTAAATGTTTCAACACTCGTACCAAGAGTA 840
DB      79278  CTATCGAAGTTGATCTCAAGACTCGGCTTAAATGTTTCAACACTCGTACCAAGAGTA 79219
QY      841  AACCGGTTTGATCCCGGATTTATTCGGAACACCGGTAATCAACCGGTAATTTGCTCGT 900
DB      79218  AACCGGTTTGATCCCGGATTTATTCGGAACACCGGTAATCAACCGGTAATTTGCTCGT 79159

```

ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes Magnoliopsida; eudicotyledons; core eudicots; Iosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis
REFERENCE	1 (bases 1 to 1605)
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Shim,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Saitou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE	Arabidopsis ORF clones
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 1605)
AUTHORS	Kim,C.J., Chen,H., Cheuk,R., Shim,P., Bowser,L., Carninci,P., Dale,J.M., Hayashizaki,Y., Ishida,J., Jones,T., Kamiya,A., Karlin-Neumann,G., Kawai,J., Lam,B., Lin,J., Miranda,M., Narusaka,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Sakurai,T., Saitou,M., Seki,M., Southwick,A., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Shinozaki,K., Davis,R.W., Theologis,A. and Ecker,J.R.
TITLE	Direct Submission
JOURNAL	Submitted (13-JUN-2003) Salk Institute Genomic Analysis Laboratory (SIGAL), Plant Biology Laboratory, The Salk Institute for Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
COMMENT	RKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA : 'RKEN Arabidopsis Full-length cDNA') : Seki,M., Narusaka,M., Ishida,J., Saitou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J., Hayashizaki,Y. and Shinozaki,K.
TITLE	The Salk, Stanford, PGEC (SSP) Consortium members constructed and sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim,C.J., Chen,H., Cheuk,R., Shim,P., Bowser,L., Chan,M.M., Chang,C.M., Dale,J.M., Jones,T., Karlin-Neumann,G., Lam,B., Lee,J.M., Lin,J., Miranda,M., Nguyen,M., Onodera,C.S., Palm,C.J., Quach,H.L., Southwick,A., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yamada,K., Yu,G., Davis,R.W., Theologis,A., and Ecker,J.R.
FEATURES	Kim,C.J. (SSP/salk) and Seki,M. (RIKEN GSC) contributed equally to this work. Shinozaki,K. (RIKEN GSC) and Ecker,J.R. (SSP/salk) contributed equally to this work as PIs.
SOURCE	Location/Qualifiers 1..1605 /organism="Arabidopsis thaliana" /mol_type="mRNA" /db_xref="taxon:3702" /chromosome="3" /clone="U16612" /note="This clone is in pUNI 51 ecotype: Columbia" 1..1605 /note="cytochrome p450 (CYP78A9)" /codon_start=1 /product="At3g61880" /protein_id="AAP6830.1" /db_xref="GI:31711948"
CDS	/translation="MTRKPTSHLAAASKSLTQNTNLATLVASIASLASLSIF WSHGSPAWGKFTLRRRTTVIPGPGLFPVSGMSMTSTLAHRCIAARERPRER LMASLGERTIVTCNDVAKEILNSVPEDRPVKESAYSLMRPAIGAPYCYRT LRKASHNEISPKOIKSETORVIANQIVCKLTROSNTGCLCPARDULTASLNMM CSVGKEVELEEEBSESSELRELVEBGZDLGLTNMDHLPMISEPDPOIRKSNUL VPKNRYNRITISDRBOTSDPEDFDVLISDGPKLSDPIIALVMIEFRGTDT VAYLIENTIAMVTADPIOSTVHEHLQOIYGRSAVESDVSHLYTAAVKEVRLAH PGELLSMARLATIDTTIIDGGRVAGCTMAWMMAIHDDHVNMPEEFERVAKEK GEVPSVARALDPLAFSGSRGVDCGNLGLTYTFMTALLHFELTSDSEKTVLL SEKRISCENMPAAALCRPRSFSV"
BASE COUNT	378 a 400 C 395 G 432 T
ORIGIN	61.3%; Score 976; DB 8; Length 1605;
Query Match	

Best Local Similarity 78.0%; Pred. No. 1.5e-285;
Matches 1238; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCTTAATCTTTGGCCCTTTTGTCCAAATGCAAGCGTTTAAAGCCAAACCAACTTGGC 78
DB 22 AGTAGCTTACTTTGGCCCTCTGTCCAAATGAGCCTCTTACTCAAAACCAATCTTGGC 81
QY 79 TTTCTCCTCCGCGCCGTCACAAATCATCTGCTGCCATATCTCTTCTATGAGCACT 138
DB 82 CTCTCTCTCCGAGCTCCCTAGCTTCTCTGCTCTTCTCTCTCTCTCTCTCTCTCTCT 141
QY 139 CCGGAGGACCTGCTGGGGGAAATACCTCTTGGCCGGTTAAATCCGGTTCAACAA 198
DB 142 CCGGAGGACCGCGATGGGGAAAGTACTTCTCCACCGC-----CGCGCT 186
QY 199 ACCGAAACGTTATTCCTCCGCTCCAAAGCCTTCCCTTGGTTGAAAGCACTCACTCAG 258
DB 187 CAACCAACCGGTATACCGGCGCAAGAGCTTACCTTTTGTCCAGACATGCTCTCAG 246
QY 259 TCAAGCACTAGCTCAACGACGATCGTGTATGACAGTGAATTCGAGCCAAAGAG 318
DB 247 TCAAACTCTGCTCAACGCTTACGCTAGCCGCAACCGCAAGAAATTTAGCGGAAAG 306
QY 319 CTCAAGCTTCACTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 378
DB 307 TTAATGGCGTTAGTTGGGAGAACTCGGTGATCTGACAGTGCATCTGATGAGCT 366
QY 379 AAAGAGATTCGAATAGCCGGTTTGTGTGATGACGCGGTTAAAGATCGGCTTACTCA 438
DB 367 AAAGAGATTCGAATAGCCGGTTTGTGTGATGACGCGGTTAAAGATCGGCTTACTCA 426
QY 439 CTGATGTTTAAACAGAGCAATGTTTTCACACAGCGTGTTAATCGGCAAGCGCTTCC 498
DB 427 CTCAATGTTTAAACAGAGCAATGTTTTCACACAGCGTGTTAATCGGCAAGCGCTTCC 486
QY 499 CGTATCGTTCGAACCACTCTTATGATCAAAACAATCAGAAAGCGGAGAGAGAGAG 558
DB 487 AAATTCGCGTATTCATCTTTTCAAGCCGAAACAGTTAAACGTTCCGAAACGAGAGA 546
QY 559 CGAGTATCTCAAGCAAGATGTTGATTTCTTGAATAACAGATAGTAAAGAA---CCC 615
DB 547 AGCGTATCGGATTCGAATCGTGAAGTGTCTCAAAACAGATAGTAAACAGATAGTAA 606
QY 616 TGTCTTGTCTGAGTGTCTTAAACAGCGCGCTGCTTAAACAGATAGTAAACAGATAG 675
DB 607 TGTCTTGTCTGAGTGTCTTAAACAGCGCGCTGCTTAAACAGATAGTAAACAGATAG 666
QY 676 GAGCAAGATGAGCTTGAATAAAACATG-----TTGAGTTACGTTGAATGCTC 726
DB 667 GGAAGAGATGAGCTTGAATAAAACATG-----TTGAGTTACGTTGAATGCTC 726
QY 727 GAAAGAGTATGATTTGCTGGAACGTTGATGAGTGAATGATCACTCTCTTGGCTATCG 786
DB 727 GAAAGAGTATGATTTGCTGGAACGTTGATGAGTGAATGATCACTCTCTTGGCTATCG 786
QY 787 GAGTTGATTCCTCAAGAGACTCGGCTGATGATTCACACTCGTACCAAGATGAAACCGG 846
DB 787 GAGTTGATTCCTCAAGAGACTCGGCTGATGATTCACACTCGTACCAAGATGAAACCGG 846
QY 847 TTTGATCCCGGATTAATCCGACACCGGATTAACACCGGATTTGCTCTGATTTCC 906
DB 847 TTTGATCCCGGATTAATCCGACACCGGATTAACACCGGATTTGCTCTGATTTCC 906
QY 907 GTCCAGCTTTGCTCTCCCTCCATGCTTCAAGTAAATATCCGACCGGCAATATGCTCC 966
DB 907 GTCCAGCTTTGCTCTCCCTCCATGCTTCAAGTAAATATCCGACCGGCAATATGCTCC 966
QY 967 GTTCTTTGGAGATGATTTCAAGAGAAACAGACAGTTCGGCTTATGATGAGTGCATC 1026
DB 967 GTTCTTTGGAGATGATTTCAAGAGAAACAGACAGTTCGGCTTATGATGAGTGCATC 1026
QY 1027 CTCCTAGAGATGCTCTTCAATCCAGTATGCAAGTCAAGGTTCAAAAGAGGTGATCAA 1086

DB 1027 CTTCGATGATGCTCTTCAATCCAGTATTCATTCAGACGTTCAATGATGATTCAA 1086
QY 1087 GTAGCCGGGAATCAAGAGCCCTAGATGAATCTGACTTGGCTCACTTCCATATCTAACG 1146
DB 1087 ATCGGGAGCATCAAGGCGCTGTGAGAGCTGACGTGCTGATATATCTAACG 1146
QY 1147 GCTGTGTGAAGAAGTATGAGCTTCATCTTCAAGCCCACTTCTATCATGAGCCCGCT 1206
DB 1147 GCTGTGTGAAGAAGTATGAGCTTCATCTTCAAGCCCACTTCTATCATGAGCCCGCT 1206
QY 1207 TTGGGCAATACAGACAGATGCTGTTGAGTGTGCTGTTTCCGGAGGAGACCAAGCAATG 1266
DB 1207 TTGGGCAATACAGACAGATGCTGTTGAGTGTGCTGTTTCCGGAGGAGACCAAGCAATG 1266
QY 1267 GTGAACATGTGGGCGGTATTCGATGATTCACACGCTGTGGTTGATCCTTTGAGTTTAA 1326
DB 1267 GTGAACATGTGGGCGGTATTCGATGATTCACACGCTGTGGAGAAATCCGTTGAGTTTAA 1326
QY 1327 CCTGAGAGTTCGTGGCAAAAGAGTGAAGTGAAGTTCGTTCTTGGGTCGATTTG 1386
DB 1327 CCTGAGAGTTCGTGGCAAAAGAGTGAAGTGAAGTTCGTTCTTGGGTCGATTTG 1386
QY 1387 AGACTTGACCTTTCGGGTCGGGTCGTGCGATTTGCCCCGGGAAAGATCTTGGTTTACT 1446
DB 1387 AGACTTGACCTTTCGGGTCGGGTCGTGCGATTTGCCCCGGGAAAGATCTTGGTTTACT 1446
QY 1447 ACCGTATGTTTGAACGCGCATGATGTTTATCATGATTTGAATG---GGGACCGCGCAT 1503
DB 1447 ACCGTATGTTTGAACGCGCATGATGTTTATCATGATTTGAATG---GGGACCGCGCAT 1503
QY 1504 ACCGTATGTTTGAACGCGCATGATGTTTATCATGATTTGAATG---GGGACCGCGCAT 1563
DB 1504 ACCGTATGTTTGAACGCGCATGATGTTTATCATGATTTGAATG---GGGACCGCGCAT 1563
QY 1507 GAGAAAGCCGTTGACCTTGTCCAGAAACGAGAGCTGTGTGATGATGCTCTCTT 1566
DB 1507 GAGAAAGCCGTTGACCTTGTCCAGAAACGAGAGCTGTGTGATGATGCTCTCTT 1566
QY 1564 CCTGCTAAATTCGCGCGGTAGGCGCAT 1591
DB 1564 CCTGCTAAATTCGCGCGGTAGGCGCAT 1591
QY 1567 GCTGCTAAATTCGCGCGGTAGGCGCAT 1594
DB 1567 GCTGCTAAATTCGCGCGGTAGGCGCAT 1594

RESULT 4
AY062748
LOCUS
DEFINITION
Arabidopsis thaliana cytochrome p450 (CYP78A9) (F21F14.5) mRNA,
complete cds.
ACCESSION
AY062748
VERSION
AY062748.1 GI:17065143
KEYWORDS
F1L1 CDNA.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 (bases 1 to 1824)
Southwick,A., Karlin-Neumann,G., Nguyen,M., Lam,B., Miranda,M.,
Palm,C.O., Bower,J., Jones,T., Bann,J., Carninci,P., Chen,H.,
Chen,K., Chung,M.K., Hayashizaki,Y., Ishida,J., Kamiya,A.,
Kawai,J., Kim,C., Lin,J., Liu,S.X., Narusaka,M., Pham,P.K.,
Sakano,H., Sakurai,T., Satou,M., Seki,M., Shim,P., Yamada,K.,
Shinozaki,K., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submision
Submitted (14-NOV-2001) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
e-mail for correspondence: arab@sequence.stanford.edu
Riken Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
Arabidopsis Full-length cDNA"): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PEBC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Nguyen,M.,

Southwick, A., Karlin-Neumann, G., Lam, B., Miranda, M., Palm, C.J.,
 Bowser, L., Jones, T., Banu, J., Chen, H., Cheuk, R., Chung, M.K.,
 Kim, C., Lin, J., Liu, S.X., Pham, P.K., Sakano, H., Shim, P.,
 Yamada, K., Becker, U., Theologis, A. and Davis, R.M.

Southwick, A., (SSP/Stanford) and Seki, M. (RIKEN GSC) contributed
 equally to this work. Shimozaki, K. (RIKEN GSC) and Davis, R.M.
 (SSP/Stanford) contributed equally to this work as PIs.

FEATURES

source

1. 1824
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="111"
 /clone="RAF109-06-018"
 /note="This clone is in pluescript
 ecotype: Columbia"

gene

CDS

1. 1824
 /gene="F21F14.5"
 54. 1658
 /gene="F21F14.5"
 /codon_start=1
 /product="cytochrome p450 (CYP78A9)"
 /protein_id="AL32826.1"
 /db_xref="GI:17065344"

BASE COUNT 477 a 438 c 427 g 482 t
 ORIGIN

Query Match 61.3%; Score 976; DB 8; Length 1824;
 Best Local Similarity 78.0%; Pred. No. 1.5e-285;
 Matches 1238; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

19 AGCTCTTAATCTTGGCCCTTTGTCCTCAATGACGGTCTAAGCCAAACCACTGGC 78
 75 AGTAGCTTACTTTGGCCCTTGTCTCAATGAGCTCTTACCAACCAATCTTGT 134
 79 TTCTCCCTCTCGCCGTCACATCATCTGGCTCGCAATATCTCTCTTTATGACCTAT 138
 135 CTCTCTCTCTCGAGGCTCCCTAGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCT 194
 139 CCCGATGACCTGTTGGGGAATACTCTTCGCGCGGTATATATCCGTTATACAA 198
 195 CCCGAGAGACCGGATGGGAAAGTACTCTCTCAACGC-----GGCGGT 239
 199 ACCGGAACCTTATCCCGGTCCAAAGGCTTCCCTTTGGTTGAAAGCATGTCATC 258
 240 CAACCAACCTGATACCGGAGCAAGAGGCTTACCTTTGTGGAAGCATGTCATC 299
 259 TCAAGCATCTAGTCAACGACGATGCTGATGATGATGATGATGATGATGATGATG 318
 300 TCAACACTCTGGCTACCGTTGATAGCGGCAACCGCAAGAAATTTAGAACCGA 359
 319 CTATGGCTTCTAGCTTACGAGAGATCTCGGTATCTGATGATGATGATGATGATG 378
 360 TTAATGCGTTAGTTAGTTGAGAAATCTCGGTATCTGATGATGATGATGATGATG 419
 379 AAAGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 438
 420 AAAGAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 479
 439 CTGATGTTTAAACAGAGAAATGTTGTTGACACACAGGTTTATCTGGGAAACGCT 498
 480 CTCAATGTTAAACCGGTATGAGTTTGGCTCTTACGAGGCTTATCTGGGAAACCT 539

499 CGATGCTTGAACCATCTCTTATGATCAAAAGAAATGAGAGCGGAGCAACGA 558
 540 AAATGCGCTTATGATCTTTTACGCCGAAACGATTTAACTTCGAAACGACAGA 599
 559 CGAGTATCTAAGCCAGATGTTGATTTCTTGAATAACAGATGATGATGATGATG 615
 600 AGCGTATGCGAATCAATATGATGATGATGATGATGATGATGATGATGATGATG 659
 616 TGTGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 675
 660 TGTGTTGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 719
 676 GAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 726
 720 GGAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 779
 727 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 786
 780 GAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 839
 787 GAGTTGATCTCAAGACTCGGATCTGATGATGATGATGATGATGATGATGATG 846
 840 GAATTTGATCTCAAGACTCGGATCTGATGATGATGATGATGATGATGATGATG 899
 847 TTTGATCTCGGATTTATCCGACACGATGATGATGATGATGATGATGATGATG 906
 900 TTTGATACCGATTTATCTGACACCGATGATGATGATGATGATGATGATGATG 959
 907 GTGACGTTTCTCTCTCCCTCATGTTTCAATGATGATGATGATGATGATGATG 966
 960 GTTGAAGTTTCTCTCTCTCATGATGATGATGATGATGATGATGATGATGATG 1019
 967 GTTCTTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
 1020 GTTCTATGAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1079
 1027 CTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
 1080 CTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1139
 1087 GTAGCGGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146
 1140 ATGTGAGACATCAAGGCTGTGAAAGTGTGATGATGATGATGATGATGATG 1199
 1147 GCTGTGTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1206
 1200 GCTGTGTAAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1259
 1207 TTGCGCATTAACAGACATGATGATGATGATGATGATGATGATGATGATGATG 1266
 1260 TTAGCATCAAGACATGATGATGATGATGATGATGATGATGATGATGATGATG 1319
 1267 GTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
 1320 GTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1379
 1327 CTGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1386
 1380 CCCGAACTTTTGTAGCCAGAAAGTATGATGATGATGATGATGATGATGATG 1439
 1387 AGACTTGACCTTTGGGCTGGGTGTGATGATGATGATGATGATGATGATGATG 1446
 1440 AGGCTTGACCTGTGGGCTGGGTGTGATGATGATGATGATGATGATGATGATG 1499
 1447 ACCGTTATTTTGAACGCGATGATGATGATGATGATGATGATGATGATGATG 1503
 1500 ACCGTTATTTTGAACGCGATGATGATGATGATGATGATGATGATGATGATG 1559
 1504 GTTAAGGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1563
 1560 GAGAGACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1619

QY 1564 CCTGCTAATTTGGCCGCTAGCGCAGTT 1591
 DB 1620 GCTGCTAATTTAGCCGCCAGCGCAGTT 1647

RESULT 5
 AB036059
 LOCUS
 DEFINITION AB036059 1902 bp mRNA linear PLN 23-DEC-1999
 Arabidopsis thaliana CYP78A9 mRNA for cytochrome P450, complete cds.

ACCESSION AB036059
 VERSION AB036059.1 GI:6633988
 KEYWORDS cytochrome P450;
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1902)
 Ico,T.

REFERENCE 1
 AUTHORS Ico,T.
 TITLE cytochrome P450 (CYP78A9)
 JOURNAL Published Only in DataBase (1999)
 REFERENCE 2 (bases 1 to 1902)
 AUTHORS Ico,T.
 TITLE Direct Submission
 JOURNAL Submitted (21-DEC-1999) Toshiro Ico, California Institute of
 Technology, Division of Biology, 1200 E. California Blvd, Pasadena,
 CA 91106, USA [E-mail: tico@cco.caltech.edu, Tel:626-395-4936]
 location/Qualifiers

FEATURES
 source
 1..1902
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /cultivar="Landsberg erecta"
 /db_xref="taxon:3702"
 /chromosome="3"
 /map="83 cM"
 /clone="28-5"
 1..1902
 /gene="CYP78A9"
 69..1673
 /gene="CYP78A9"
 /codon_start=1
 /product="cytochrome P450"
 /protein_id="PAA8569.1"
 /db_xref="GI:6633989"
 /translation="MARTKLDTSLLALSKSLITQTNLALSLVVASLASLISLP
 WSHGGRAMKQPIHRRQTTVIGPRGLFVPGSMISMTLHRCIAARAEPRER
 LMAPSLGSTRIVTCNPDAKEILNSVFADRPKESAYSLMNRRAIGFAPGVWRT
 LKRLASNLHSPKQIKRSEORVIANQIVKLTQKSTNGLCFARDLKTASLNNM
 CSVPKSEVLEEEHEVSEBELVEBVDILGLNMTDHPMLSEDPORIKRNLNM
 VPKNRFEVNRILSDHREQTRDSPDFVLLSLGPKLSDPIILAVMEMIRGNTD
 VAVLIEWILAMVLPDIQSTVHNELDQIVGRSAVESDVSLVYLAVKEVRLH
 PPGILSMARLITDITIDGRVAPGTAVNMWMLAHDPHWMENLEPEEFVAKB
 GEVEFSYLGSLIRLAPPGSGRNRGNLGLTYTTTATLILHEFMTLSDKTYDL
 SEKURLSCENMPLAKLRPRRSV"

BASE COUNT 507 a 448 c 439 g 508 t

ORIGIN
 Query Match 61.3% Score 976; DB 8; Length 1902;
 Best local similarity 76.0%; Pred. No. 1.5e-265;
 Matches 1238; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY 199 ACCGAAACGTTATTCGGGTCGAAAGGCTTCCCTTGTGGAAGCATCTCACTATG 258
 DB 255 CAACACACCGTGAATCCGGGCGCAAGAGGCTTACCTTTGTGGAAAGCATCTCTATG 314
 QY 259 TCAAGCATCTAGCTCAACGACGAAATCGCTGATCGATGGAATTCGAGCAAGG 318
 DB 315 TCAACACTCTGGCTCACCGCTTGCAATACCGCAACCGACGAATTTAGAGCGGAAG 374
 QY 319 CTGATGCTTTACAGCTTAGAGAGACTCGCTGATCGTCACTCAATCCGACGTACG 378
 DB 375 TTAATGGCGTTAGTTAGTGGAGAACTCGCTGATCGTCACTCAATCCGATGAGCT 434
 QY 379 AAAGGATTCGATAGCGCGGTTTTCGATGACCGGTTAAGAAATGGCTTACTCA 438
 DB 435 AAAGGATTCGATAGCGCGGTTTTCGATGACCGGTTAAGAAATGGCTTACTCA 494
 QY 439 CTGATGCTTTACAGAGCAATGCTTTTGCACACAGCGGTTTACTGCGAACGCTTCG 498
 DB 455 CTGATGCTTTACAGAGCAATGCTTTTGCCTTACAGCGGTTTACTGCGAACGCTTCG 554
 QY 459 CGTATCGCTTGAACCATCTCTTTAGTACAAACAAATCAGAAAGCCGACACGACGA 558
 DB 555 AAATATCGGCTTAATCATCTTTTCAGCCCGAAACGATTAACGTTCCGAAACGACGAGA 614
 QY 559 CGAGTATCTCAAGCCAGATGAGTTGCTTGAAGAAACAGAGTATGACGAA---CCC 615
 DB 615 ACCTGATCGCAATCAATCTGTAAGTGTCTCAAAACAGAGTACACCAAGGCTTC 674
 QY 616 TGTTTTGTTCGTGAGTCTTAAACGCGCTGCTTAAACATGATGTCTCTGTATTC 675
 DB 675 TGTTCGACAGCTGATGATCAAAACGCGCTTAAACATGATGTCTCTGTATTC 734
 QY 676 GGACAGAGTATGAGCTTGAAGAAACCATG-----TTGATTCGTTGAATGGTC 726
 DB 735 GGAAAGAAATACAGCTTGAAGAAACCATGAGTATGAAAGAGTATGATGATGCTG 794
 QY 727 GAAGAGGTTATGATTTGCTCGGAACGTTGATGATGATGATGATGATGATGATG 786
 DB 795 GAAGAGGTTATGATTTGCTCGGATGATGATGATGATGATGATGATGATGATG 854
 QY 787 GAGTTGATCTCAAGAGCTCGGCTGATGATGATGATGATGATGATGATGATGATG 846
 DB 855 GAGTTGATCTCAAGAGCTCGGCTGATGATGATGATGATGATGATGATGATGATG 914
 QY 847 TTTGATCCCGGATATATCCGACACCGTATCAACCGGATGATGATGATGATGATG 906
 DB 915 TTTGATCCCGGATATATCTGACACCGTATCAACCGGATGATGATGATGATGATG 974
 QY 907 GTGACGTTTGTCTCTCCCTCATGCTTCAATGATGATGATGATGATGATGATGATG 966
 DB 975 GTTACGTTTGTCTCTCTCTCATGCTTCAATGATGATGATGATGATGATGATGATG 1034
 QY 967 GTTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
 DB 1035 GTTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
 QY 1027 CTGCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
 DB 1095 CTGCTAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
 QY 1087 GTAGTGGGAATCAAGGCGCTTCAATGATGATGATGATGATGATGATGATGATG 1146
 DB 1155 ATGTGGGACATCAAGGCGCTTCAATGATGATGATGATGATGATGATGATGATG 1214
 QY 1147 GCTGTGTGAAGAGATATGAGCTTCACTTCAAGGCGCTTCAATGATGATGATG 1206
 DB 1215 GCTGTGTGAAGAGATATGAGCTTCACTTCAAGGCGCTTCAATGATGATGATGATG 1274
 QY 1207 TTGGCATACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266
 DB 1275 TTAGCATACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334

QY 1267 GTGAACATGTGGGCGGATTCGATCATCAGCGTGGGTTGATCCTTTGAGTTTAA 1326
 Db 1335 GTGAACATGTGGGCGGATTCGATCATCAGCGTGGGTTGATCCTTTGAGTTTAA 1394
 QY 1327 CCGTGAAGTTCGTTGGCAAAAGAGTGAAGGAGTTTGGTTCTTGGGTCGATTTG 1386
 Db 1395 CCGGACGTTTGTAGCCAGAGAGTGAAGTGAAGTTCTCGTTCTTGGGTCGATTTG 1454
 QY 1387 AGACTTGACCTTTGGGTCGGGTCGGTGGATTTGGCCCGGAGAAATCTTGGTTTAACT 1446
 Db 1455 AGGCTTGACCGTTGGGTCGGGTCGGTGGATTTGGCCCGGAGAAATCTTGGTTTAACT 1514
 QY 1447 ACCGTTAGCTTTGACCGGCGATGATGTTACATGAGTTGAATG---GGACCGTCGAT 1503
 Db 1515 ACCGTTAGCTTTGACCGGCGATGATGTTACATGAGTTGAATGAGTACCGTCGAT 1574
 QY 1504 GGTAAACGGCTTGAATTCCTGAGAACTGAGGCTTTCTTGGAGATGCTTAATCTTCTT 1563
 Db 1575 GAGAAACCGTTGACTTGTCCGAAACTGAGGCTCTCGTGAAGATGCTTAATCTTCTT 1634
 QY 1564 CTTGCTTAATTCGCGCGCTTGGGCGGCTT 1591
 Db 1635 GCTGCTTAATTCAGCCCGGCGGCGGCTT 1662

RESULT 6
 LOCUS AY086928 1878 bp mRNA linear PLN 14-APR-2003
 DEFINITION Arabidopsis thaliana clone 29661 mRNA, complete sequence.
 ACCESSION AY086928.1 GI:21405652
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 1878)
 Haas,B.J., Volfovsky,N., Town,C.D., Troupkhan,M., Alexandrov,N.,
 Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
 Full-length messenger RNA sequences greatly improve genome
 annotation

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA
 This clone sequence is one of 5,000 Ceres full-length cDNAs made
 available to TIGR and Genbank. The following quality assessment of
 this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (ATG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the WS or LAR ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 Genest carried out the library production and sequencing of the
 full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.
 Location/Qualifiers

source 1.1878
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="29661"
 51.1655
 /codon_start=1
 /product="cytochrome P450 (CYP78A9)"
 /protein_id="AA064492.1"
 /db_xref="GI:21592543"
 /translation="MATKIDTSLLALSKSLTQTNLALSLVSLASLALSLF
 WSPGPGWAGKFLHRRQTVIPGRGLPEVGSMSLMSNTLAHRCIALATKFAER
 LMAFSLGETRVLVTCNPVAKELINSPEVADPVEKSAVSLMFNAIOPAPGVWRT
 LKIASNHLFSPKQIKSRSEKSVIANQVILKQSNYKGLCPARDILKIASLNMM
 CSVQKREYELIEEHSEVSELRELVEGDLIGTLMTDLPLSEFDPQIRSCNL
 VPKVRFPVRLITISDRRECRDSPPDPVYLISDPPDLSDPDIIVALEMTEVRLH
 VALLLEMLARVHLDPDQSTYHNLDOVGRSRAVRSADVSLVYLAVAYEVRLH
 PGPPLSLMARLAIITDILLIDGRVPEGTAAVMWALADPHYEMPLEKPERPVAKS
 GEVSEVLSGLRIAPFPGSRVPCGKMLGLTIVFWATLHLHEEXLTPSDEKIVDL
 SEKRLRSCMANPLAKLPRRSFV"
 BASE COUNT 501 a 441 c 436 g 499 t 1 others
 ORIGIN

Query Match 61.2%; Score 975.6; DB 8; Length 1878;
 Best Local Similarity 77.9%; Pred. No. 2e-285;
 Matches 1237; Conservative 1; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCCATATCTTCCCTTGTGTCATATCAGCGTCTTAAGCCAAACCACTTGC 78
 Db 72 AGTACCTTACTTTGGCCCTCTTCCAAATGAGCTCCTTAACCAATCTTGTCT 131
 QY 79 TTCTCCCTCTGCGCCCTCAACATCATCTGAGTCCCAATCTCTCTTTATGACCTAT 138
 Db 132 CTCTCTCTCTCTTACCTCCCTCCAGCTTCTCTGCTCTTCTCTCTCTCTCTCTCT 191
 QY 139 CCGGATGACCTGCTGGGGGAAATACCTCTTGGCGGCTTAATATCCGTTATCAAA 198
 Db 192 CCGGATGACCTGCTGGGGGAAATACCTCTTGGCGGCTTAATATCCGTTATCAAA 236
 QY 199 ACCGAAACCTTATCCCGGTCGAAAGGCTTCCCTTGGTGAAGCATGTCTCATG 258
 Db 237 CAACACCGCTGATACCCGGGCCAAGAGGCTTACTTTGTGGAAGCATGTCTCATG 296
 QY 259 TCAAGCATCTTACCTACCGAAGATCGGTGATGAGCTGAGAAATTCGAGCCAGAG 318
 Db 297 TCAACACCTTGTGCTACCGTTCATGACCGAAGCATGAGAAATTCGAGCCAGAG 356
 QY 319 CTGATGCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTACCTTAC 378
 Db 357 TTATGCGCTTATGTTGGGAAAGATCGGTGATGCTGCACTGCAATCCTGATGAGCT 416
 QY 379 AAGAGATTTGGAATAGCCCGGTTTGTGTCATGACCGGTTAAAGATCGGTTACTA 438
 Db 417 AAGAGATTTGGAATAGCCCGGTTTGTGTCATGACCGGTTAAAGATCGGTTACTA 476
 QY 439 CTGATGTTTAAACAGAGCAATTTGTTTGAACACAGGTTTACTTGGCGAAGCTTGGC 498
 Db 477 CTGATGTTTAAACAGAGCAATTTGTTTGAACACAGGTTTACTTGGCGAAGCTTGGC 536
 QY 499 CGTATCGTTGGAACATCTCTTTAGTCAAAACCAATCGAAGCCGAGAGCAAGCA 558
 Db 537 AATATCGGCTTATCATCTTTTCAAGCCGAAACAGTTAAAGTTTCGAAAGCAAGCA 596
 QY 559 CGATGATCTCAACGAGATGTTGATTTCTTGAAGAAACAGATGATACGA---CCC 615
 Db 597 AGCGTATCGGATCAAAATCGTGAAGTGTCTCAAAAACAGATGATACGAAGGCTC 656
 QY 616 TGTTTTGTGTGATGTTCTTAAACGGCGTCTTAAACATGATGCTCTGATTTTC 675
 Db 657 TGTTTTGTGTGATGTTCTTAAACGGCGTCTTAAACATGATGCTCTGATTTTC 716
 QY 676 GACACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 726

Db 717 GGAAGAAATACAGAGCTTGAGAGAGCATGAAGAAAGTAGAGCTGACGTGAATTGGTC 776
 Qy 727 GAAGAAGATTATGATTGGCTCGAAGCTGGAATTGAGTGAATCACTTCCTTGGCTATCG 786
 Db 777 GAAGAAGATTATGATTGGCTCGAAGCTGGAATTGAGTGAATCACTTCCTTGGCTATCG 836
 Qy 787 GAGTTGATCTCTCAAGAGCTCCGGTCTAGATGTTCCACACTCGTACCAAGGTAACCGG 846
 Db 837 GAATTGATCTCTCAAGAGCTCCGGTCTAGATGTTCCACACTCGTACCAAGGTAACCGG 896
 Qy 847 TTTGGTATCCCGGATTATATCCGAACACCGTAATCAACCGGTAATTGCTCCGTGATTTTC 906
 Db 897 TTTTGAACCGGATTAATCTGTGACACCGGTAACCAAACTGTACTACCGAGTGACTTC 956
 Qy 907 GTGACGTTTGGCTCTCCCTCCATGTTCCAGTAATTAATCCGACCCGACATATACGCC 966
 Db 957 GTTGACGTAATGCTCTCTCCATGTTCCAGTAATTAATCCGACCCGACATATACGCC 1016
 Qy 967 GTTCTTGGAGATGATTAATCAGAGAGACAGACAGTTCGGTCTTAATCGAGTGAATC 1026
 Db 1017 GTTCTATGGAGATGATTAATCAGAGAGACAGACAGTTCGGTCTTAATCGAGTGAATC 1076
 Qy 1027 CTCGCTAGAGATGATCTCTTCAATCCAGATATGCAATCAACGGTACAAAAGAGCTGATCA 1086
 Db 1077 CTGCTAGAGATGATCTCTTCAATCCAGATATCAATCGAGTTCACATGAGCTGATCA 1136
 Qy 1087 GTATCCGGGAATTAAGAGCCCTAGATGATGATCTTGGCTCACTCATATCAATCAAG 1146
 Db 1137 ATCTGAGAGCATTAAGAGCTGTTGAAGAGCTGAGCGTGTCTAGTAATCTTACG 1196
 Qy 1147 GCTGTGAGAGAGATTAAGAGCTTCAATCTTCCAGCCCACTTCAATCATGAGCCCGT 1206
 Db 1197 GCTGTGAGAGAGATTAAGAGCTTCAATCTTCCAGCCCACTTCAATCATGAGCCCGT 1256
 Qy 1207 TTGGCCATACAGACAGATGTTGATGATGTTGTTTCCGAGAGGACCAACAGCAATG 1266
 Db 1257 TTACCAATACAGACAGATGTTGATGATGTTGTTTCCGAGAGGACCAACAGCAATG 1316
 Qy 1267 GTGAACATGAGGCGGTATCGATGATCAACAGCTGTGAGTGAATCTTGGAGTTTAA 1326
 Db 1317 GTGAACATGAGGCGGTATCGATGATCAACAGCTGTGAGTGAATCTTGGAGTTTAA 1376
 Qy 1327 CCTGAGAGTTCGTGGCAAAAGAGTGAAGTTCGTTTGGTTCGGTGGATTTG 1386
 Db 1377 CCCGAAAGTTCGTGGCAAAAGAGTGAAGTTCGTTTGGTTCGGTGGATTTG 1436
 Qy 1387 AGACTTGACCTTCCGGTGGGTGCTGAGATTTGCCCGGGAAGATCTTGGTTTACT 1446
 Db 1437 AGCTTGACCTTCCGGTGGGTGCTGAGATTTGCCCGGGAAGATCTTGGTTTACT 1496
 Qy 1447 ACCGTTATGTTTGGACGGCATGATGTTACATGATTTGAATG--GGACCGTCCGAT 1503
 Db 1497 ACCGTTATGTTTGGACGGCATGATGTTACATGATTTGAATGATCCCGTCCGAT 1556
 Qy 1504 GGTACCGGCGTTGATTAATCTGAGAACTGAGAGCTTCTTCCGAGATGAGTATCTT 1563
 Db 1557 GAGAAAGACCGTTGATTTCCGAGAACTGAGAGCTTCTTCCGAGATGAGTATCTT 1616
 Qy 1564 CCGTCTAAATGCGCGGTAGCGGCGAGTT 1591
 Db 1617 GCTGCTAAATTAATGCGCGGCGGCGAGTT 1644

RESULT 7
 APT21F14/C
 LOCUS APT21F14 142001 bp DNA linear PLN 02-FEB-2000
 DEFINITION Arabidopsis thaliana DNA chromosome 3, BAC clone F21F14.
 ACCESSION AL138642
 VERSION AL138642.1 GI:6899881
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatoophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; euroside II; Brassicales; Brassicaceae; Arabidopsis.
 1
 Choisme, N., Robert, C., Broctier, P., Wincker, P., Cateolico, L.,
 Arciguenave, F., Saurin, W., Weissenbach, J., Mewes, H.W., Lemcke, K.,
 Mayer, K.F.X., Quetier, F. and Salanoubat, M.
 Unpublished
 2 (bases 1 to 142001)
 EU Arabidopsis sequencing project.
 Direct Submission
 Submitted (02-FEB-2000) MIPS, at the Max-Planck-Institut fuer
 Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail:
 lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project
 Coordinator: Marcel Salanoubat and Francis Quetier, Groupement
 d'Interet Public, Centre National de Sequencage - GENOSCOPE, 2 rue
 Gaston Cremieux, BP191, 91006 Evry Cedex, France;
 http://www.genoscope.cns.fr
 Information on performance of analysis and a more detailed
 annotation of this entry and other sequences of chromosomes 3, 4
 and 5 can be viewed at: <http://www.mips.biochem.mpg.de/Proj/thal/>.
 FEATURES
 source
 location/Qualifiers
 1..142001
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /variety="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 1..56180
 /note="overlap to BAC F15G16, please refer also to Acc.
 no. EMBL:AL132959 for analysis and annotation"
 5113..55887
 /gene="F21F14.10"
 join(5113..55695,55835..55887)
 /gene="F21F14.10"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="CB71891.1"
 /db_xref="gi:6899882"
 /translation="WVIVKRCITFDNYIYLHCLLOTNNHKKOKEEDVVKTKKK
 KDTSPDADNETMILITSLQHLVLAQRSRLQYSSPVITQVSLSLMSGDDSLRA
 SLPFSEBEPGEPKHVHNPSPYSKLDLPPEMLPEDETKRPTANRHYRVLAGEKWP
 RKERKDDAVYVAGSGSYIPSEDDTKTRIKVPMKKLKKASVPA"
 5113..55695
 /gene="F21F14.10"
 /number=1
 55696..55834
 /gene="F21F14.10"
 /number=1
 55835..55887
 /gene="F21F14.10"
 /number=2
 57967..57993
 /note="27 bp TCT tandem repeat"
 58167..59822
 /gene="F21F14.20"
 join(58167..58190,58872..59376,59461..59822)
 /gene="F21F14.20"
 /codon_start=1
 /product="transcription factor BBFa"
 /protein_id="CB71892.1"
 /db_xref="gi:6899883"
 /translation="MDATRWQGFQEMINVKEMQMSITNNTPQOFTPTATNTP
 NATASNGSGGANTNNYATWETRKARPOEKNPCRNSTTRCYNNYSLTQPRYCK
 GGRVYTBEGSLRNVPVGGSSRKRSSTPLASPSNPKLPDLNPILESQIPNKNK
 DNLISLSPWODHHHALLRLRNGVSRGKMTPLFGQWMDNSVLSYLSGPTWPDY
 KQSNNTLSSTIHQICIGHTINSNRADNDNDMGARVLPFSMDKELSLSTQEK
 SHQNTYKMGFSNTGSSW"
 58167..58190
 /gene="F21F14.20"
 /number=1
 58191..58871
 /gene="F21F14.20"
 /number=1

exon

intron

repeat_unit

gene

CDS

misc_feature

gene

CDS

exon

intron

exon

Db 70256 TTAATGGCGTTAGTTGGAGAACTCGCGTATCGTACAGTGAATCGTATGAGT 70197
 Qy 379 AAAGAGATTCGAATAGCCGCTTTTCTGATGACCGGTAAAGATCGGCTTACTCA 438
 Db 70196 AAAGAGATTCGAATAGCCGCTTTTCTGATGACCGGTAAAGATCGGCTTACTCC 70137
 Qy 439 CTGATGTTTAAAGAGCAATGTTTTCACACAGGCTTTTCTGCGCAACGCTTCCG 498
 Db 70136 CTGATGTTTAAAGAGCAATGTTTTCACACAGGCTTTTCTGCGCAACGCTTCCG 70077
 Qy 499 CGTATCGCTGCAACCACTCTCTTTGATCAAAACAAACAGAAAGCCAGACGCA 558
 Db 70076 AAAATCGGCTTAAACATCTTTTCAGCCGAAACAGATTAACCTTCCGAAACGACAGA 70017
 Qy 559 CGAGTATCTCAGCCAGATGTTGATGTTTCTTGA AAAACAGAGTATGTAAGAA ---CCC 615
 Db 70016 AAGCGATCGCAATCAATCGTGAAGTCTCAAAAACAGAGTAAACCAAGGCTC 69957
 Qy 616 TGTGTTGTTGCGATGCTTAAACCGCGCTGCTTAAACAGATGCTCTGATTC 675
 Db 69956 TGTGTTGCGATGCTTAAACCGCGCTGCTTAAACAGATGCTCTGATTC 65897
 Qy 676 GCACAGAGTATGAGCTTGA AAAACAG ---TTGAGTTAGCTGAATGCTC 726
 Db 69836 GGAAGAGATACGAGCTTGAAGAGAGCATGAAGAGTGAAGTACGTAATGCTG 65837
 Qy 727 GAAAGAGTATGATTTGCTGGAACGTTGAATGGAATGCTGCTCTGCTGCTATCG 786
 Db 69836 GAAAGAGTATGATTTGCTGGAACGTTGAATGGAATGCTGCTCTGCTGCTATCG 69777
 Qy 787 GAGTTGATCTCTCAAGAGCTCGGCTTGAATGTTCAACTGCTACCAAGGTAACCG 846
 Db 69776 GAAATTTGATCTCTCAAGAGCTCGGCTTGAATGTTCAACTGCTACCAAGGTAACCG 69717
 Qy 847 TTTGATATCCCGATATATCGGAACCGGTAATGGAATGCTGCTCTGCTGCTATTC 966
 Db 69716 TTTGGAACCGGATTTATCTCTGACCAAGTGAACCAATCGTCACTCAAGAGTATTC 69657
 Qy 907 GTGACGTTTGTCTCTCTCTCCATGTTTCAATTAATTCAGACCGGACATATCGCC 966
 Db 69656 GTTGAAGTATGCTCTCTCTGAGTGTCTGATTAATTAATTCAGACCGGATTCATCGCC 69597
 Qy 967 GTTCTTTG----- 974
 Db 69536 GTTCTATGAGTATGATGACAGTGTCTGATTTTACTTTGTAAGAAAACTTAC 69537
 Qy 975 ----- 974
 Db 69536 AAAAATCATATTTATGTAAGTATGATGATTTAATTAATTAATTAATTAATTTGAT 69477
 Qy 975 ----- 974
 Db 69476 AATTAAATATGCAATTTTAAAGTATTAAGTATTAAGTATTTGGCTTTTAAATTAACA 69417
 Qy 975 GGAAGTATGATTTCAAGAGACAGACAGATGTCGGCTTAAATGAGTATTCCTCTAG 1034
 Db 69416 GGAATATGATTTCAAGAGACAGACAGTGTCTTTTGAATGAGTATTCCTCTAG 69357
 Qy 1035 GATGCTCTTCAATCAAGATATGCAATCAACGTTCAAAAGAGTGAATCAAGTACG 1094
 Db 69356 GATGCTCTTCAATCAAGATATGCAATCAACGTTCAAAAGAGTGAATCAAGTACG 69297
 Qy 1095 GAAATCAAGAGCCCTTGAATGATTTGATTTGCTTCACTTCCATATTAACGCTGTGT 1154
 Db 69296 AAGATCAAGAGCCCTTGAAGAGTGTGAAGTGTCTCTTAATATTAACGCTGTGT 69237
 Qy 1155 GAAAGAGTATTTGAAGCTTCACTCTCAAGCCCACTTCAATCAAGGCGCGTTGGCAAT 1214
 Db 69236 GAAAGAGTCTTGAAGCTTCACTCTCAAGCCCACTTCAATCAAGGCGCGTTGGCAAT 69177
 Qy 1215 AAGAGAGATCGTTGATGCTCTTGTCTTCCGCAAGGACACAGCAATGTTGAACAT 1274
 Db 69176 CACAGAGATCGATCGAGGCTGTGTCTTCCGCGGAGACACCCCAATGTTGAACAT 69117

Qy 1275 GTGGCCGATATGCAATATCCACAGCTGTGGCTTGAATCTTTGAGATTTAACTGAGAG 1334
 Db 69116 GTGGCCGATATGCAATATCCACAGCTGTGGAGAAATCTTTGAGATTTAACTGAGAG 69057
 Qy 1335 GTTCGTGCAAAAGAGTGAAGTGAAGTGTTCGTTCTTGGTGGATTTGAAGTTCG 1394
 Db 69056 TTTGTAGCAAGAGAGTGAAGTGAAGTGTTCGTTCTTGGTGGATTTGAAGTTCG 68997
 Qy 1395 ACCTTGCGGCTGT 1454
 Db 68996 ACCTTGCGGCTGT 68937
 Qy 1455 GTTTTGAAGCGAGTATGTTACATGATTTGAATG ---GGAGCGTCCGATGTAACG 1511
 Db 68936 GTTTTGAAGCGAGTATGTTACATGATTTGAATG ---GGAGCGTCCGATGTAACG 68877
 Qy 1512 CGTTGACTATCTAGAAACCTGAGGCTTCTTCCGAGATGAGTAACTCTTCTGCTAA 1571
 Db 68876 CGTTGACTATCTAGAAACCTGAGGCTTCTTCCGAGATGAGTAACTCTTCTGCTAA 68817
 Qy 1572 ATTGGCCGTAAGGCGAGTT 1591
 Db 68816 ATTACGCCCGAGCGCAGTT 68797
 RESULT 8
 AF022463
 LOCUS AF022463 1611 bp mRNA linear PLN 02-MAR-1999
 DEFINITION Glycine max cytochrome P450 monooxygenase CYP78A3p (CYP78A3) mRNA,
 complete cds.
 ACCESSION AF022463
 VERSION AF022463.1 GI:2739007
 KEYWORDS
 SOURCE Glycine max (soybean)
 ORGANISM Glycine max (soybean);
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
 Glycine.
 REFERENCE 1 (bases 1 to 1611)
 AUTHORS Simlinszky,B., Corbin,F.T., Ward,E.R., Fleischmann,T.J. and
 Dewey,R.E.
 TITLE Expression of a soybean cytochrome P450 monooxygenase cDNA in yeast
 and tobacco enhances the metabolism of phenylurea herbicides
 JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (4), 1750-1755 (1999)
 MEDLINE 99145622
 PUBMED 9990096
 REFERENCE 2 (bases 1 to 1611)
 AUTHORS Simlinszky,B., Dewey,R.E. and Corbin,F.T.
 TITLE Direct Submission
 JOURNAL Submitted (04-SEP-1997) Crop Science, North Carolina State
 University, Box 7620, Raleigh, NC 27695, USA
 FEATURES
 source
 1. 1611
 /organism="Glycine max"
 /mol_type="mRNA"
 /db_xref="taxon:3847"
 1. 1611
 /gene="CYP78A3"
 20. 1591
 /note="CYP78A3"
 /gene="cytochrome P450 monooxygenase"
 /note="cytochrome P450 monooxygenase"
 /codon_start=1
 /product="CYP78A3p"
 /protein_id="A894592.1"
 /db_xref="GI:2739008"
 /translation="MTSHIDNNIMILATSKCTOENLAVLIMSLMTMPYKSH
 PGEPAMKRTYTPPLSLITPGRKFPPLISGKIMTSLAHRRIAPAAATPRAKIAPRS
 LGDRIVLVCHPVAKEILISVFAUPRVESYSIMLRRAIGFASYSLSLRKIA
 SNHLFCEPQIKASELQSOIAQWVHILNKRRLSRVQVAKSLSNMWSVFQGE
 YKLHPNSGMEDELIVDQGYDLGLFNNAADHLPLAHPDAQNIIRRCNNLVPMVNR
 VGIIAHSRASKETNRDFFVDVLLSLPEPDQSLSDSMIAVLVEMIFRGDTVAVLIV

ILAMALPHYQSKVSEILDAVNGKARAVAEADVAVMTYLPAVKEVILRHPPELIS
 WARSINDTIDGTHVPAQTATTAATTAICADPHVMDPLEPMPERBTRGDAEESI
 LGSDFRLAPGSRACGKTLGMATVNFVASHLHEFEWPSDEKVDULEVLKLS
 EMANPLTVKXPRRG"

BASE COUNT 369 a 457 c 407 g 378 t
 ORIGIN

Query Match 40.9%; Score 651.6; DB 8; Length 1611;
 Best Local Similarity 66.5%; Pred. No. 1.1e-186;
 Matches 1025; Conservative 0; Mismatches 484; Indels 33; Gaps 5;

QY 61 AGCCAAACCAACCTGCTCTCCCTCCGCGCTCAACATCATCTGCTGCCATCT 120
 DB 74 ACCCAAGAAACCTTGACATGGGTCTTTGATCAAGGCTCACTGGTTAACATGACT 133
 QY 121 CTCTCTTATGACCTATCCGCTGACCTGTTGGGGAAATACCTCTTGGCGGTTA 180
 DB 134 TTCTATTACTGCTACACCCGCTGCTCTGCTGGGCAAGTACTACAC-----TA 186
 QY 181 ATATCCGCTTCATACAAACCGAAACGTTATCCCGCTCCAAAGGCTTCCCTTGGTT 240
 DB 187 CTCTCCCCCTTTGAA-----TCAATCCGCTCCCAAGGCTTCCCTTATT 235
 QY 241 GGAAGCATCTACTCATGTCAAGCACTAGCTCACGACGATGCTGATGACGCTGAG 300
 DB 236 GGAAGCATGCGCTCATGACTTCC--CTGGCCCATCACCGTATCGACGCGCGCC 292
 QY 301 AAATTCGAGCAAGAGGCTCATGGCTTCACTTAGAGAGACTCGGCTGATGCTACG 360
 DB 293 ACATGACAGCCCAAGGCTCATGGCTTATGCTCGGCAACACGCTGATGCTACG 352
 QY 361 TCGAATCCCGACGTAGCGAAGAGATTCGAATAGCCCGCTTTTCTGATGACGCTT 420
 DB 353 TCGAATCCCGACGTAGCGAAGAGATTCGAATAGCCCGCTTTTCTGATGACGCTT 412
 QY 421 AAGAGATCGCTTATCTCATGTGATTAAGAGCAATGTTTGGACACACAGGCTT 480
 DB 413 AAGAGATCGCTTATCTCATGTGATTAAGAGCAATGTTTGGACACACAGGCTT 472
 QY 481 TACTGGCGAAGCTTGGCGGTATCGCTTGAACCATCTTTTATGACAAACAAATCGA 540
 DB 473 TACTGGCGAAGCTTGGCGGTATCGCTTGAACCATCTTTTATGACAAACAAATCGA 532
 QY 541 AAGAGCGAAGCGAAGAGATGATCTCAAGCCGATGATGATGATTTCTGAACAAACAG 600
 DB 533 GCGCTGAGCTCCAAAGCTCTCAATAGCGCGCCCAATGATGATCTCT--AAATGAC 589
 QY 601 AGTAGTAACGAACCCGTTTGTGCTGAGTTGCTTAAACCGCGCTGCTTAAACATG 660
 DB 590 AACCGCAACCGAGCTTAGGTGCTGCAAGTGTGAAAAAGCTTCTGCTGATGATG 649
 QY 661 ATGTGCTGTATTCGACAAAGATGAGCTTGAACAAACCA-----TGTGAG 711
 DB 650 ATGTGCTGTATTCGACAAAGATGAGCTTGAACAAACCA-----TGTGAG 709
 QY 712 TTAGTGAATGCTGACAAAGGATGATGATGCTGGAACGTTGAATGATGATGATGAC 771
 DB 710 CTGGAATATTAATGACCAAGTATGACCTGTTGGGCTGTTTAAATTTGGGCGACAC 769
 QY 772 CTCTCTTGGCTATGAGATTTGATCTCAAAAGCTCGCTGATGATGATGATGATGATG 831
 DB 770 CTCTCTTGGCTATGAGATTTGATCTCAAAAGCTCGCTGATGATGATGATGATGATG 829
 QY 832 CCAAGGTAACCGGTTGATCCGATATATCGAACAACCGTATGATGATGATGATGATG 891
 DB 830 CCAAGGTAACCGGTTGATCCGATATATCGAACAACCGTATGATGATGATGATGATG 889
 QY 892 TTGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 951
 DB 890 ACNATGCTGATTTGTTGAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 949
 QY 952 CCGGACATATCGCGCTTCTTGGGAGATGATGATGAGGACAGTGTGGGCTC 1011

DB 950 TCCGACATGATGCTGATCTTGGGAAATGATATTCAGAGGACGACGCTGACGCTT 1009
 QY 1012 TTAATCGAGTGAATCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
 DB 1010 TTGATGAGTGAATCTGCTGAGATGATGATGATGATGATGATGATGATGATGATGAT 1069
 QY 1072 AACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
 DB 1070 GAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
 QY 1132 CTTCATATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1191
 DB 1130 ATGACGATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1189
 QY 1192 CTATCATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
 DB 1190 CTATCATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
 QY 1252 GGAACACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
 DB 1250 GGAACACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309
 QY 1312 CTTTGAAGTTTAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
 DB 1310 CCACTGATTTATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
 QY 1372 CTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1431
 DB 1370 CTGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
 QY 1432 AATCTGATTTTACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491
 DB 1430 ACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1489
 QY 1492 GGAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551
 DB 1490 GGAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1549
 QY 1552 GCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1593
 DB 1550 GCTAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591

RESULT 9
 ATH552847/c
 Locus
 DEFINITION
 Arabidopsis thaliana T-DNA flanking sequence, left border, clone 345B11.
 ACCESSION
 AJ552847
 VERSION
 AJ552847.1
 KEYWORDS
 left border; T-DNA flanking sequence.
 SOURCE
 Arabidopsis thaliana (thale cress)
 ORGANISM
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosid II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE
 1. Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Craud, C., Derose, R., Pelletier, G., Lepoint, L., Caboche, M., and Lecharny, A.
 T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
 EMBO Rep. 3 (12), 1152-1157 (2002)
 TITLE
 JOURNAL
 MEDLINE
 PUBMED
 22363535
 12446565
 2 (bases 1 to 740)
 Balzerque, S.
 Direct Submission
 Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE
 PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment (b) resulting from the PCR were directly sequenced from the left or the right border

CDS

AGSVNORNSGSYNTYSEYDSANHGQOFENSNIMQOQPLQGSFNPLEYDPANH
GGWMLSDYIDLQOQVPIYLAPEYSESMIMKIVIDENEPFLVDBERTSMQOHSHRPK
PVSGLVLPDSDSTGSMIFEDTSSDTSVSSDEPGHTRIDIPSLNIEPLHANYA
OEPKROSKKIVISSOKCEWKAABDSIKIPSTNTVQOSMIVLENAQNNYKNNMI
GVLEFISWITLIIVG"

complement (join(6617..6953,7266..7351,7464..7603,
7916..7998,8087..8166,8273..8368))

/note="hypothetical protein"

/codon_start=1

/evidence=not experimental

/product="T25K16.2"

/protein_id="AAF26477.1"

/db_xref="GI:6715650"

CDS

/translation="MAASERHCVGCGPRVKSFLFOYSPGNIRLMKCNCKEVADEYE
CERNVCNHFPLSLFGPKYRHRVLTALNPATVNIQVKNFVSTRCVGEIHRDTIK
SPBLIDRSLLRKSDSESSFSDSPVLSTIKVLIGVLSANAFITISALITKGLNEV
SRESILLQVWERPMSVIFFDILLTNSMALKQOTRMEQMIVPCCCYFSQCKF
VEKPMTESTMTRCIAVCLAHILIRFLVGGQIFEEPTIFLIQGSLLQYMSYFRIV"

complement (11566..12642)

/note="putative RAB.8 protein gi|3695373"

/codon_start=1

/evidence=not experimental

/product="T25K16.3"

/protein_id="AAF26476.1"

/db_xref="GI:6715649"

/translation="MDLSAPPTTSSDQEDRDQOELTNSNIGASSSGSPGNNNLM
NMIPPEKEHMFDPKVTBPDVQKLNRLVITPROHARYEPPLDSSNNNGTLLIPQDRG
KMMRRRYSYNMSQSYVVTCKMSRPRVEKKIKIDAGDIYSFGQIGDESRSPKLYDWH
RPMDSLVQAOQNGFNFPPTISQISNRTHPEINSEVPIHGLNIGNORSTYNT
ORQEFVGYGYNLAGRCYTGSPLDHNRIVSEBPLVIDSVFVPGRLTPVLPPLDP
PSTACKRLRLFGVNMECNDYNOQDESWLPRGEIGASSSSSALRLNLSIDHDDND
DDDDDDDOFAFKGKSLSLNFP"

join(123221..24174,24244..24357,24412..24664,24743..25137,
25226..25445,25527..25711,25783..25905,25994..26478,
26564..26730,26814..26983,27074..27235,27320..27415,
27505..28133,28314..28507,28592..28782,28862..30013,
30112..30518,30604..30781)

/note="similar to UFDI protein emb|CAB10321.1; similar to
ESTs gb|H36434..gb|A1996152.1"

/codon_start=1

/evidence=not experimental

/product="T25K16.4"

/protein_id="AAF26461.1"

/db_xref="GI:6715634"

/translation="WMDEPREATIKSYWMDACEDISCDLIDLVSEFPSSAVAN
ESTDNGVINDPFGGIDHILDSIKNGGGLPNNGVSDTNSQINVEVTPQVIAKTIVE
NLOKNGKGRDESKKEGDKRKARVCSYSESRNLGSRGVNNSREGDFNNKRT
NMWDAGNNKKRKECNRYRDRGDRREVGRMDVGSNELVYRSGTEADHERDYK
VSGGRCEDYKAEKRSKREKREKVEBOARVOLDVBOAKANTTIAFETGAGCT
LIALILIKSHEDILMSQNRKXLSVPLVEKVPVYVQVPPKKGQAEVITNQCPOVGH
CEMGQDFWDSRRKORERESKOFPLKTSFPLSSIVQVWTRQIILNLRSSITIMEI
IDILLIDECNAAVKGHPESLVVSEBFTTPKDKRAIFGMAFPVNLGVSQVDAK
KIRLNETKLDSTVCTIKORLEKEVPMSEIIVEYDKAATMSLHETIKQMLAVE
AAQASRSKQMGPMGARADGADDELROYVSERTESDAANLHKLALIVYTTAEIG
QMCAYVQGSFSLQOSDERVNFQVDKROESYLSVAVLSLQCELLGGAALKEAAVE
GPEGNONADMEBERGELPDDPVVSGEHEVDYGAVALDGVGTQVSGILITLKYOH
TADFRALYVERVVAALVYPKRIKYPALBLSITRCSXNIGHNSQMSQMDOT
ISFRDGHVTLVAITSVAEGDILQCNVWFDLAKVLAITQSRGAKRPSGTYIL
IVERYIKSFKNYILIFVTHQIISTDMSCVTCRKNVSHAALRNARNSSEETLRKEAI
BESTDLSHLDKSTRLLSIDAVPGIVKVAITGAMVLSNVSQVJHVFYCSQILQSRVAIL
REBSWEKEKEGATEYSCRLQPCMAPELIEGVSMSMAAOQVNIYSACKLH
BKGAFTDMLLPKSGGDAEKADODDEPGFGRARHEFVEBGAADVILKQWWSGK
KVCESKFLHLMVYVNRVCDPSSKODPESEVSEFALLFGENLIDLEVYSMSDILVAV
AMITKSLAKGSLDTENOLSLIKKHFYRMSVLDVDSVSTTTPMPARAIVFVY
TNTSKRPIKGINMELVEKITTTTANDPLOPARPVLTGHTERTLQGDREYGGKL
RNHYVGGKSHPTVYIGRVAASFDVVRAGLIPVDAEKEVEEDLSKQMMADQGM
VVEDLIGKIVTAHSGKRFYVDSICYDMSAETSFRKSGVLTPLNRYVADYKQKIX
VQDRLEFYFHNLRLLRLYKSSIMLFRYGVLDNCKQOPLIKRGVSCYNLSPR
FEGSGSEVLDKTYVFLPELGVVHPSGLLIGAGQALPSIMRVSMLLAVALCN
LSYPISTKILTEALTAASCOETFEYERLELDGAYLKAVSRFLFLKPKQHEGOLT
RMOQVMSNMVYOPALVKGLOSIOADFRAPSRMSAPGVPEVPEIDGGSFDE
EOKPYSEMSDVPEDEGMDGELGDLSSRYLSSKTLDADVEALIGYVYVGGKIAA
NHLKNIIGIHVEDBPDVSGITLKNVAVPSVLKSIDPGLLEALIKTEREKGLVEAI

CDS

THASPPSGVGCYQRLREVGDVLDHLITRLHFTTYTTLPPGRLLDLRAAVNNENFA
RVAVRKRLHLYRHSSSLKLEKOVNITKQDSIFSKFKCLTWMLFVQITBFRKEVO
TSSSRGRFSLGDCAKPKVGLDIVESIAQILFDSKDTTAARKVFOPLQPVTP
ETLPMHAYLEOERCOQOAGELEYASRSQNTATVEVITDVOVGAONPKQAQKL
AARNALALKEKEIAESKEKILNNAGEDQGENENKNGHQFTRQTLNDICLR
NMPSYTCVKEGGAHAKRFPGVAVNTSRGWTDECIGRPMSPVKAQSAVALIL
ELNMTFS"

complement (join(33084..31126,31223..31304,31341..31515,
31635..31700,31790..31897,31984..32049,32133..32161,
32249..32372))

/note="putative inorganic pyrophosphatase gi|3510259;
similar to ESTs gb|T42316..gb|A1994042.1..gb|A1994013.1,
emb|229202"

/codon_start=1

/evidence=not experimental

/product="T25K16.5"

/protein_id="AAF26475.1"

/db_xref="GI:6715648"

/translation="MSEETKDQRLQRPAPRLNERILSLSRSVAAHPWDELIGPG
AFOIFNVVVEIKRSGKVKYELDKTGLIKVDILYSVYVPHNYGFVRLTCEMDPI
DYLVMOEVLPGCFRLPRAIGMLMPINDQEKDILIAVCVDPEYKATYDIKEAPPH
RSEIRRPREDICLFLQCSLSIFSIDLSTNKNKEVAVANDFLPSBAVAEALQYMSD
LYAEYILHLTR"

complement (join(33694..34029,34103..35173,35269..35349,
35432..35701,36326..36387,36512..36623,36725..36763))

/note="putative late elongated hypocotyl emb|CAA07004;
similar to ESTs gb|A1993521.1..gb|A650979"

/codon_start=1

/evidence=not experimental

/product="T25K16.6"

/protein_id="AAF26474.1"

/db_xref="GI:6715647"

/translation="MDNTSGEELAKARPYITIKORRWTEDEHERFLEALRYGR
AMORIEHIGTIAVDIRSHAKPFTFKRGSFMTFQLEKEAEVKGIPIVQALDIB
IIPPRKRKPTVPRKPGNNGTSOVSADALVSASSQSLADFLDEKSPFS
EKTSRKENQDNCQGVSTVNYKPLPTKYSGLIENSTYVNAVODVYKKNQDNT
EYTSKQNTYFMHTRADIVYNIKACIPNHSQVMSQDFMFAPEHFGNANTOAT
TASATTAHQAPACHSODDYRSFQJOSTISNSLWSTLONPAAHAAATPAAVMP
YASVSGSGSTSPSSPSTIAAATAAATAVMAWASGLIPVCAPELVTPSTV
AVTPAMTEMDVETNQPPEKONTALODNTAAKSPASSSDSDTGTGYNLADSKTN
DKIEEVVTAVHDSNTAKKLVDRSSGSGTSGSAETDALDKMKEDYET
DENQPVITLNNRKIKRDNNSNNATITPMSKEVEBEGIAFOALFAEBRLPQSFSP
QVAVENNRQSDTSMPLAIPFKSODSCADDSGVVIMIGVCTKSLKTRGTGKPYKRC
SHEVKSQVGNINQSDDEKCYKRLRREGAS7"

complement (join(38600..38756,38838..38989,39111..39516,
39915..40031,40377..40579))

/note="similar to Medicago truncatula MtN2 gi|3193308;
similar to EST gb|H77065"

/codon_start=1

/evidence=not experimental

/product="T25K16.7"

/protein_id="AAF26473.1"

/db_xref="GI:6715646"

/translation="MAGDMQGRVAVKESPVIVMWSNVAKGSVNALVKKALDVGVRH

CDS

complement (join(33694..34029,34103..35173,35269..35349,
35432..35701,36326..36387,36512..36623,36725..36763))

/note="putative late elongated hypocotyl emb|CAA07004;
similar to ESTs gb|A1993521.1..gb|A650979"

/codon_start=1

/evidence=not experimental

/product="T25K16.6"

/protein_id="AAF26474.1"

/db_xref="GI:6715647"

/translation="MDNTSGEELAKARPYITIKORRWTEDEHERFLEALRYGR
AMORIEHIGTIAVDIRSHAKPFTFKRGSFMTFQLEKEAEVKGIPIVQALDIB
IIPPRKRKPTVPRKPGNNGTSOVSADALVSASSQSLADFLDEKSPFS
EKTSRKENQDNCQGVSTVNYKPLPTKYSGLIENSTYVNAVODVYKKNQDNT
EYTSKQNTYFMHTRADIVYNIKACIPNHSQVMSQDFMFAPEHFGNANTOAT
TASATTAHQAPACHSODDYRSFQJOSTISNSLWSTLONPAAHAAATPAAVMP
YASVSGSGSTSPSSPSTIAAATAAATAVMAWASGLIPVCAPELVTPSTV
AVTPAMTEMDVETNQPPEKONTALODNTAAKSPASSSDSDTGTGYNLADSKTN
DKIEEVVTAVHDSNTAKKLVDRSSGSGTSGSAETDALDKMKEDYET
DENQPVITLNNRKIKRDNNSNNATITPMSKEVEBEGIAFOALFAEBRLPQSFSP
QVAVENNRQSDTSMPLAIPFKSODSCADDSGVVIMIGVCTKSLKTRGTGKPYKRC
SHEVKSQVGNINQSDDEKCYKRLRREGAS7"

complement (join(38600..38756,38838..38989,39111..39516,
39915..40031,40377..40579))

/note="similar to Medicago truncatula MtN2 gi|3193308;
similar to EST gb|H77065"

/codon_start=1

/evidence=not experimental

/product="T25K16.7"

/protein_id="AAF26473.1"

/db_xref="GI:6715646"

/translation="MAGDMQGRVAVKESPVIVMWSNVAKGSVNALVKKALDVGVRH

CDS

complement (join(33694..34029,34103..35173,35269..35349,
35432..35701,36326..36387,36512..36623,36725..36763))

/note="putative late elongated hypocotyl emb|CAA07004;
similar to ESTs gb|A1993521.1..gb|A650979"

/codon_start=1

/evidence=not experimental

/product="T25K16.6"

/protein_id="AAF26474.1"

/db_xref="GI:6715647"

/translation="MDNTSGEELAKARPYITIKORRWTEDEHERFLEALRYGR
AMORIEHIGTIAVDIRSHAKPFTFKRGSFMTFQLEKEAEVKGIPIVQALDIB
IIPPRKRKPTVPRKPGNNGTSOVSADALVSASSQSLADFLDEKSPFS
EKTSRKENQDNCQGVSTVNYKPLPTKYSGLIENSTYVNAVODVYKKNQDNT
EYTSKQNTYFMHTRADIVYNIKACIPNHSQVMSQDFMFAPEHFGNANTOAT
TASATTAHQAPACHSODDYRSFQJOSTISNSLWSTLONPAAHAAATPAAVMP
YASVSGSGSTSPSSPSTIAAATAAATAVMAWASGLIPVCAPELVTPSTV
AVTPAMTEMDVETNQPPEKONTALODNTAAKSPASSSDSDTGTGYNLADSKTN
DKIEEVVTAVHDSNTAKKLVDRSSGSGTSGSAETDALDKMKEDYET
DENQPVITLNNRKIKRDNNSNNATITPMSKEVEBEGIAFOALFAEBRLPQSFSP
QVAVENNRQSDTSMPLAIPFKSODSCADDSGVVIMIGVCTKSLKTRGTGKPYKRC
SHEVKSQVGNINQSDDEKCYKRLRREGAS7"

complement (join(38600..38756,38838..38989,39111..39516,
39915..40031,40377..40579))

/note="similar to Medicago truncatula MtN2 gi|3193308;
similar to EST gb|H77065"

/codon_start=1

/evidence=not experimental

/product="T25K16.7"

/protein_id="AAF26473.1"

/db_xref="GI:6715646"

/translation="MAGDMQGRVAVKESPVIVMWSNVAKGSVNALVKKALDVGVRH

Query Match 29.84; Score 474.4; DB 8; Length 86436;
Best Local Similarity 59.64; Pred. No. 1,7e-132;
Matches 1087; Conservative 0; Mismatches 486; Indels 251; Gaps 7;

Q#	Q	D#	D	Score	Length	Identical	Mismatches	Indels	Gaps
Q1	ATGGCTACGAACCTCGAAGCTCCTTAATCTTTGCTCCCTTTTGTCCAAATGACAGCTTCTA	D60	ATGGCTACGAACCTCGAAGCTCCTTAATCTTTGCTCCCTTTTGTCCAAATGACAGCTTCTA	60	100	0	0	0	0
Q2	ATGGAAGCCCAATAGAAATTTTGCGGTATTTGCTTTCGCAACAAATTCATATTATTAC	D84507	ATGGAAGCCCAATAGAAATTTTGCGGTATTTGCTTTCGCAACAAATTCATATTATTAC	84507	100	0	0	0	0
Q3	AGCCAAACCAACCTGCTTCCTCCTGCTGCGCGTACAAATCATCTGCGCATATCT	D120	AGCCAAACCAACCTGCTTCCTCCTGCTGCGCGTACAAATCATCTGCGCATATCT	120	100	0	0	0	0
Q4	ATGCAACCAACATTTTGCTTCTCTCTGCTGCGCATATCTGCTGCTTATCAATAC	D84447	ATGCAACCAACATTTTGCTTCTCTCTGCTGCGCATATCTGCTGCTTATCAATAC	84447	100	0	0	0	0
Q5	CTCTTCTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTC	D177	CTCTTCTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTC	177	100	0	0	0	0
Q6	ATCGTATTTTGGTCTATCCGGGTGACCGCGCTGGGGAATAATCTTCACTCGCGG	D84387	ATCGTATTTTGGTCTATCCGGGTGACCGCGCTGGGGAATAATCTTCACTCGCGG	84387	100	0	0	0	0
Q7	TTATATCCGGTTCATAC-----AAACCGGAACGTTATCTCCGGTCCAAAGGCTTC	D231	TTATATCCGGTTCATAC-----AAACCGGAACGTTATCTCCGGTCCAAAGGCTTC	231	100	0	0	0	0

Db 84386 TTATATCTCTCGATTACACCGAATAATCTATCCCGGCTTAGAGGTT 84327
 QY 232 CCTTTGGTGGAGAGATGTCATCTATGTCAGACATCTAGCTCAGCGAATCGCTGAT 291
 Db 84326 CCACTTGTGGAGAGATGAGCTTAGGCTCAGCGCAAGTGCTCATAGCGCATAGGCT 84267
 QY 292 GCAGCTGAGAAATTCGAGCCGAGAGAGGCTCATGGCTTTCAAGCTTAGAGAGACTGCGGTG 351
 Db 84266 GTGGCTGAGATGAGTAAAGCGCAAGCGGCTCATGGCTTTAGGCTCGATGATCTAAGGTTG 84207
 QY 352 ATGTGACCGTGCATCCCGAGCGTAAAGAGAGATCTGAAATAGCCCGGTTTGTGCTAT 411
 Db 84206 GTGGTACGCTGCTCTCTGCGGTGGCAAGAGATCTAAGAGCTTCCGTTTGTGCTAG 84147
 QY 412 CGACCGGTTAAGATTCGCTTCTACCTGATGTTTAAAGAGCAATGGTTTGCACCA 471
 Db 84146 CGACCGGTTAAGCAACCGCTTACGCTTATGATGTTTAAAGAGCAATGGGATTTGCTCC 84087
 QY 472 CACGCTTTTACTGCGCAAGCTTCCGCTTACGCTTCCGACCATCTCTTTAGTACAA 531
 Db 84086 AATGCTACTTATGGCGTACGCTGCGCTGCTAGGCTGCAACATCTTTTAAAGCCGAG 84027
 QY 532 CAATTCAGAGAGCGGAGAGCAAGCAAGAGATCTCAAGCGAGATGGTTGAGTTCTT 591
 Db 84026 CAATTCAGAGATCGAGAGATCGAGAGCGGCTGATGCGACTCAGATGGTGAATGCGATT 83967
 QY 592 GAAAAACAGAGTACTTAAAGAACCTCTGTTTGTCTGAGATGCTTAAAGAGGCGTGCCT 651
 Db 83966 G---CAGTAAACCTTAATCCGCTGCGAGTGGATGCTTAAAGAGGCGTGCCT 83910
 QY 652 AAGAACAGATGCTGCTGATTCGAGCAAGAGATGAGCTTGAAGAAAAACCA----- 704
 Db 83909 TGTAACTGATGGTGTGTTTGTTCGAGAGAGATGATGATGAGTGAATTAACAATTG 82850
 QY 705 -----TGTGAGTTACGTGAATGCTGAAAGAGGTTATGATTTGCTCGAAGCTTGAAT 759
 Db 83849 GAATCTGAATGCTTAAAGAGGTTTGTGAGAGAGGAGTACATCTTCAAGTACGTTAAT 83790
 QY 760 TGGACTGATCACTTCCCTGCGCTATCGGAGTTTGAAGCTCCTGAAGACTCCGCTCTAAGATG 819
 Db 83789 TGGACCGACATCTCTCTGCTTACCGGTTTAAATTTCAACAAATCCGCTTTAGTGC 83730
 QY 820 TCCACACTCTGACCAAGAGTAAACCGGTTTGTATCCGAGTTATATCCGAACCGCTAAT 879
 Db 83729 TCGCAGCTCGTACCGAAAGTAAATCTGTTATGAGCGGTATCATATGAAACAACGCTCT 83670
 QY 880 CAAGCCGCTGATTTGCTCGCTGATTTGCTGACGTTTGTCTCCCTCCCATGGTTCAGAT 939
 Db 83669 GCCACGGGTAACTTTC-----TTACAGTAACTTCTCTTCAAGGTTCCGAA 83622
 QY 940 AATATATCCGACCGGACATATCGCCGCTCTTTG----- 974
 Db 83621 AATATATGAGATCCGACATGTTGCTGTTTGGGAGATTTTGTCTTTATCACTA 83562
 QY 975 ----- 974
 Db 83561 TTCAATCCGATATCATGATATCAACATTAAGTCAACTAGATTAACATTAAGAAATCC 83502
 QY 975 ----- 974
 Db 83501 AATAAATTTACTAATTAAGTATGATTAAGTATTAACATTAACATTAAGTATGTTGTT 83442
 QY 975 ----- 974
 Db 83441 TTAATCAATCATGAGAAATGATTAAGAAATTTCTCTACGATATTAACCTTTTGGT 83382
 QY 975 -----GGAGATGATATTCAGAGGAAAGACAGACAGTTCGCTTAAATCGAGTGCATCC 1027
 Db 83381 CCGTTTAAAGAAATGATTTTAAAGGAGAGAGACATCTTTCGCTTTTGGTGCAGTGCAG 83322
 QY 1028 TCGCTAGATGCTCTCTCATCAGATATGCAATCAACGCTTAAACAAAGAGCTGCATCAAG 1087
 Db 83321 TAGCGAGATGCTGATGATCCCAAGTTCAATTAACGCTTCAACATGAGCTTGAACGAG 83262

QY 1088 TAGTCGGAAATCAAGAGCCCTTAAATGATCTGACTTGGCTGACTTCAATCTAAGCG 1147
 Db 83261 TCGTTGGAGATCAAGAAACCGTGAATGATGAGCTTCAATCACTACATGATTAACGG 83202
 QY 1148 CTGTGTGAAAGAGATTTGAGGCTTCAATCCCTCCAGGCCCACTTCAATGAGGCCGTT 1207
 Db 83201 CTATGATCAAGAAAGTGTGAGGCTGCATCACCAGGATCACTGCTTCTTGGGACAGAC 83142
 QY 1208 TGCCCATTAACAGACAGATGCTTAAATGCTGCTTGTTCGGCAGGAGACACAGCAATGG 1267
 Db 83141 TGTCTATTAACAGACACTTCCCTAATGATATACAGTGGCGGCTGAGACACCGCATGG 83082
 QY 1268 TGAACATGAGGCGGATTCGATATCCGATATCCACAGTGGGTTGATCTTTGAGTTTAAAC 1327
 Db 83081 TCAACATGAGGCTATAGACATGACATCCACAGTGGGAGATCTTTAGATTTAAGC 83022
 QY 1328 CTGAGAGTGTGTGCAAAAGAGAGTGAAGTTTTCGCTTCTTGGTCCGATTTGA 1387
 Db 83021 CTGAGAGTGTGTGCTTAAAGAGGTTGAAGCTGAGTCTCTGTTTTCGGGCTGATCTGA 82962
 QY 1388 GACTTGCACCTTGGGCGGTGCGGTGCGATTTTGCAGGAGAAATCTTGGTTTACTA 1447
 Db 82961 GGTGGCACCGTTCGGGTGCGGTGAAGAGGTTTCCCTCGAAGAAATTTGGAGCTTACAA 82902
 QY 1448 CCGTATGTTTGAAGCGGAGATGATGTTTACATGAGTTGAATGGGAGC---GTCCGATG 1504
 Db 82901 CGGTGTGTTTGGGTTGCAACGCTCTTGCATGATTTGAGTGGCTTCCAGCGTCAAG 82842
 QY 1505 GTTACGGGCTGACTTATCTGAGAAACAGAGGCTTCTTGGAGATGCTATCTCTTC 1564
 Db 82841 CTAAACCTCCAGATCTCTCGAGGTTTGAAGGCTCTCGTGAATGCTTGTCCACTCA 82782
 QY 1565 CTGCTAAATTCGCGCGGTAGCGCA 1588
 Db 82781 TCGTTAAGCTAAGCTCAAGGGCTA 82758

RESULT 11
 AC023628/c
 LOCUS 104001 bp DNA linear PLN 03-AUG-2000
 DEFINITION Arabidopsis thaliana chromosome I BAC FEF3 genomic sequence,
 complete sequence.
 AC023628
 VERSION AC023628.3 GI:7212005
 KEYWORDS HTG.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 104001)
 2 (bases 1 to 104001)
 3 (bases 1 to 104001)
 AUTHORS
 REFERENCES
 JOURNAL
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (16-FEB-2000) DNA Sequencing and Technology Center,
 Stanford University, 855 California Avenue, Palo Alto, CA 94304,
 USA
 3 (bases 1 to 104001)

ALRSVTPNISKDKAKLGDANSYINELAKLWMEAEERELGSSNPISIDSDINV
 QTSGEDVTAIRINCPLESHPSARIIFAEESKVEVINSNLESDVLTATFVKSSELT
 KERLISALSRBQNSVOSRISGR"
 53198..54813
 /gene="F6P3.8"
 /join(53198..54103,54187..54813)
 /gene="highly similar to dbj|BAA92894.1, cytochrome P450
 from Petunia x hybrida"
 /product="putative cytochrome P450"
 /protein_id="AA097323.1"
 /db_xref="gi:9665139"
 /translation="MTILVSLFAVILINVLKRWLKASAKQRLPQPPRLPIGN
 LLOGLPRLHDLASLCKRGPPLYLRLGNDATITTDPRITRILIRODVRSRPT
 LAAYHLVYGGDVALAPMGPBMRKMRICHEHLITRLESFTORAEARLIRDPV
 KRSETPRINLKEVLGAFSMANVTMLGQFPGLSVSKAEFLHTKLFML
 GYLIGDYLPEFMRWVDPSCGEKMRDVEKVDENFKIIDEHRALEDKDKMDF
 VDLISLPGENKAMEDVEIKALIDMIAAADTSVTEMMAEAIKQPRYMKIQ
 EELDNVGSNRMDESDVHLNLCVRETFPMHAPGPELIPHSVRAATTINGYIP
 AKRVFINHGLGNTKIMDVEDFPERHMPVSGSRVLSHGPDKLIPFAGSRKX

Query Match 29.8%; Score 474.4; DB 8; Length 104001;
 Best Local Similarity 59.6%; Pred. No. 1,7e-132;
 Matches 1087; Conservative 0; Mismatches 486; Indels 251; Gaps 7;

QY 1 ATGGCTAGAACTGGAAGCTCTTAATCTTGCCCTTTGTCCAAATGAGGCTCTA 60
 Db 25772 ATGAGAACCGAATAGAAAGTTTGTGGTATTGCTCTGTCAATAAATCAATTTAC 25713
 QY 61 AGCCAAACCAACTGCTCTCTCTCTCTGCGCTCACAATCATCTGCTCCCAATCT 120
 Db 25712 ATGCAACAACATTTGCTCTCTCTCTCTGCGCTCAATCATCTGCTTACCATAC 25653
 QY 121 CTCTTTTATGACCTATCCCGGTGGAAGCTTGCGGGAAATACCTCTTC---GGCGG 177
 Db 25652 ATCGTATTTGCTCTACTCCGAGTGAACGCGCTGGGAAATACCTCTCTCACTGCGCG 25593
 QY 178 TTAATATCCGCTTCATAC-----AAAACCGAAACGTTATTCGCGTCCAAAGGCTTC 231
 Db 25592 TTTATTTCTCGATTAACAACGAATATCAATCATCTGCTTACCATAC 25533
 QY 232 CTTTGTGTTGGAAGATGTCTACTATGTCAAGCACTAGTCAACCAAGAAATGCGTAT 291
 Db 25532 CCACTTGTGGGAAGATAGGCTTATGTCAGCAACGTCGCTCATCAGCGGATAGCGTCT 25473
 QY 232 GCACTGAGAAATTCGAGCCAAAGAGCTCATGGCTTCACTTATAGAGAGACTCGCGTG 351
 Db 25472 GTGGTGAATGAGTAAGCCAAAGGCGCTCATGGCTTATAGCTCGGATATCAATAGGTG 25413
 QY 352 ATCTGTCAGTCAATCCGAGCGTAAGCAAGAGATTCGAATAGCCGCTTTTGTCTGAT 411
 Db 25412 GTGTGAGAGTCACTCTGCGTGAAGAGAGATCAATCAAGTTCGCTTTTGTCTGAC 25353
 QY 412 CGAACCGGTTAAAGATCGGCTTACTCACTGATGTTTAAACAAGCAATGTTTGGACA 471
 Db 25352 CGAACCGGTTGACGAACCGCTTACGCTTATGATGTTTAAACAAGCAATGTTTGGCTCC 25293
 QY 472 CACGGTGTCTTACGCGAAGCGTTCGCGTATCGCTTCGAACCATCTCTTATGTAACAAA 531
 Db 25292 AATGTAATCTTATGCGTACGCTGCGTGGTATAGGCTCGAAACCATCTTTTAAACCGAAG 25233
 QY 532 CAATTCAGAGAGCCGAGACGCAAGAGATGATCTCAAGCCAGATGTTGATTTCTT 591
 Db 25232 CAATTCAGATCGAGAGATGAGAGCGGATGATGAGATCTCAATGATGATGATGATGATG 25173
 QY 592 GAAAAACAGATGATGAGCAACACCTGTTTGTCTGATGTTGCTTAAACGCGCTGCTT 651
 Db 25172 G---CACGTAAACCTTAATCCGCTGTGAGATGCTGATTTGCTCAAAAACGCGCTG 25116
 QY 652 AACCAATGATGCTCTGATTTTGGACAAGAGTATGAGCTTGAAGAAAAACA----- 704
 Db 25115 TGTAAATATGAGTGTGTTTGGGAGAGATGATGATGAGTCAATTAACAACCTTG 25056

QY 705 -----TGTGAGTTACGGAAGAAATGTCGAGAGAGATATGATTTGCTCGAACCTGTAAT 759
 Db 25055 GAATCTGAATGCTTAAAGGCTTTGTTGATGAGAGAGATGATCTTCTAGATGATGATTAAT 24996
 QY 760 TGAAGTATCACTTCTCTGCTGCTATCGGAGTTGATCTCAAGATCTCCGCTCTAGATGT 819
 Db 24995 TGAGCCAGACATCTTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24936
 QY 820 TCACACTCTGTAACCAAGATTAACCGGTTGATCCCGATATATCCGACACCGTAAAT 879
 Db 24935 TCGAGCTCTGTAACCAAGATTAACCGGTTGATCCCGATATATCCGACACCGTAAAT 24876
 QY 880 CAACCGGTTGATTTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 939
 Db 24875 GGCACGGTAACTTCT-----TGACATGTAATCTTCTCTCAAGGTTCAAGAA 24828
 QY 940 AATTTATCCGACCCGAGATATATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 974
 Db 24827 AATTTATCAGATCCGACATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24768
 QY 975 ----- 974
 Db 24767 TTCAATCCGATATGATGATATCAACATATGTAATGTAATGTAATGTAATGTAATGTA 24708
 QY 975 ----- 974
 Db 24707 ATAAATTTATTAATTAAGTATGATGTAATGTAATGTAATGTAATGTAATGTAATGTA 24648
 QY 975 ----- 974
 Db 24647 TTAATCAATATGAGAAATGATTAAGAAATTTCTCTAGATATTAATCTTTTGTGT 24588
 QY 975 -----GAGATGATATTCAGAGAACAGACAGACAGTTCGCTCTTAAATGATGATGATCC 1027
 Db 24587 CCGTTAGAAATGATATTTAGGAGAACGACAGACAGTTCGCTCTTAAATGATGATGATCC 24528
 QY 1028 TCGTGTGATGCTCTCTTCAAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1087
 Db 24527 TACGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24468
 QY 1088 TAGTCGGAATCAAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1147
 Db 24467 TCGTTGAGATCAAGAGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24408
 QY 1148 CTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1207
 Db 24407 CTATGATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24348
 QY 1208 TGGCATTAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1267
 Db 24347 TGTCTTAACAGACATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24288
 QY 1268 TGAACATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1327
 Db 24287 TCAACATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24228
 QY 1328 CTGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1387
 Db 24227 CTGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24168
 QY 1448 GACTTGACCTTTGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1447
 Db 24167 GCTTGACCTTTGCGGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 24108
 QY 1448 CCGTTATGTTTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1504
 Db 24107 CCGTTATGTTTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 24048
 QY 1505 GTAACGCGTGAATTAATGAGAAATGAGGCTTTCTTGCAGATGAGGTAATCTCTTCTTC 1564
 Db 24047 CTAACTCTCAGATCTCTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 23988

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.rge

Page 20

QY 1565 CTGCTAATGGCCGTAGGCGCA 1588
DB 23987 TCGTTACGTACGTCAAGCGCTA 23964

RESULT 12
LOCUS AR070567 1799 bp DNA linear PAT 18-FEB-2000
DEFINITION Sequence 3 from patent US 5907082.
ACCESSION AR070567
VERSION AR070567.1 GI:7221455
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1799)
AUTHORS O'Neill,S. and Nadeau,J.
TITLE Ovine-specific gene expression
JOURNAL Patent: US 5907082-A 3 25-MAY-1999;
FEATURES
source 1..1799
/organism="unknown"
BASE COUNT 456 a 362 c 428 g 548 t 5 others
ORIGIN

Query Match 22.3%; Score 355; DB 6; Length 1799;
Best Local Similarity 58.0%; Pred. No. 2.8e-96;
Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTCGAGCCCAAGAGGCTCATGCTTTCAGCTTAGAGAGACTCGGCTGCTCACTGCTC 363
DB 273 TTCGAGCCCTCCGCTCATGCTTTCAGCTTAGAGAGACTCGGCTGCTCACTGCTC 332
QY 364 AATCCGAGCTAGCGGAAGAGATTTGAATAGCCGGTTTTCGCTGATCGACGGTTAA 423
DB 333 CACCCCAAAACCGCAAAAGAGATTTCTTCAAGCCCGCTTTCGCTGATCGACGGTTAA 392
QY 424 GAATGCGCTTACCTCATGATGTTTACAGAGCAATGGTTTGCACACAGGGTTTAC 483
DB 393 GAATCAGCATGCAATCTTCTGTTTATCGCGCTTAGGGTTTGCCTCATTTGGGATTC 452
QY 484 TGGCGAAGCTTTCGCGCTATCGCTTGAACCATCTCTTATGTAACAAATCGAAGA 543
DB 453 TGGAGAACTCGAGAGAGATTCGTCACATATCTTTCAGTCGCGGAGATTCATCG 512
QY 544 GCGGAGACGAAGAGAGATTCGTCACAGCCAGATGTTGAGTTCTTGAAGAAAGA-- 601
DB 513 TTCGAGAGACGAGAGATGATGTCGAGAGAGATGTCGGGATTCGAGAGAGATGATG 572
QY 602 -GTAGTACGAAACCTGTTTGTTCGTGATGTCCTTAAACGCGCTGCTTAACAACATG 660
DB 573 GAGAGAAATGAGATGTTAGAGAGAGAGATGTTCACTACGGGCTTTGAAATACATC 632
QY 661 ATGTGCTCTGATTCGAGCAAGATGATGCTTGAAGAAACCATGTTAGTTAGTGA 720
DB 633 ATGTGATGTTTTCGAGAAAGTTGATTTGCAAGAGATGAGGGTTGAGCTGAG 692
QY 721 ATGATC--GAGAGAGATGATGATTTGCTCGAAGCTTGAATGATGATCACTTCT 777
DB 693 TTGATCCTTAAGAGAGATGATGATTTGATGAGATCTTCACTCGGGGATGATCTT 752
QY 778 TGGCTATCGAGATTTGATCTCAAGATCCGCTCTAGATGTTTCCACTCTGTAACAG 837
DB 753 CTTTGGGATGTTAATTTGCAAGATGAGAGAGATGCAAGACCTTGTGGCTAAG 812
QY 838 GTAACCGGTTTATCCGATTTATCCGAAACACCGTAATCAAC----- 885
DB 813 GTCAATGATTTTGTGAAGAGATCATAGACGACATTAAGAGAGACCAACCGCGTAAG 872
QY 886 -----GATATTTGCTGATTTTGTGACGTTTGTCTCCCTCATGTTGAT 939
DB 873 ATTGATGAGGTGAAGGTGAAGATTTGTTGATGCTCTTGTGAGAGAGAAAGAT 932

QY 940 AATATCCGACCCGAGACATAATCGCGTTCTTTGGAGATGATATTCAGAGAACAG 999
DB 933 AGACTCTCAGATCTATATGATGATGATGATGATGATGATGATGATGATGATGAT 992

QY 1000 AAGTTCGCTTATTCAGATGATGATGATGATGATGATGATGATGATGATGAT 1059
DB 993 ACTGTCGATCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1052

QY 1060 TCAACGTCACAAACAGCTGATCAAGTATGTCGAGAAATCAAGAGCCCTAGATGAT 1119
DB 1053 TCGAAGGACAAATGATGATGATGATGATGATGATGATGATGATGATGAT 1112

QY 1120 GACTGCTTCACTTCCATATCTAAGCGCTGCTGCTGAGAAAGATGATGATGAT 1179
DB 1113 GATATCCAGACTTCTTATCTTCAATCTATGATGATGATGATGATGATGAT 1172

QY 1180 CAGAGCCACTTCTATCATGAGGCGCTTTCGACATACAGACAGATCGTTAGTGT 1239
DB 1173 CTTGGGCTTATGATGATGATGATGATGATGATGATGATGATGATGAT 1232

QY 1240 CTTGTCGCGCAGGACCAAGCATGATGATGATGATGATGATGATGATGAT 1299
DB 1233 ATGATTCCTGCTGAGACGACTGATGATGATGATGATGATGATGATGAT 1292

QY 1300 GTGTGGTGTATCTTGTGATGATGATGATGATGATGATGATGATGATGAT 1359
DB 1293 AACTGGCTAGGCTTAACATTAATCTGATGATGATGATGATGATGATGAT 1344

QY 1360 GAGTTTCGCTTTCGCTGATGATGATGATGATGATGATGATGATGATGAT 1419
DB 1345 ----TCAATATCTTGTGATGATGATGATGATGATGATGATGATGAT 1400

QY 1420 TGGCCCGGAGAGATCTGTTTATCTACCTGATGATGATGATGATGATGAT 1479
DB 1401 TGGCCCGGAGAGATCTGTTTATCTACCTGATGATGATGATGATGATGAT 1460

QY 1480 GAGTTTANATGGGACCGCTCGATGATGATGATGATGATGATGATGATGAT 1539
DB 1461 AGCTTCAATTTGCTCTTC---GAGAAATGATGATGATGATGATGATGAT 1517

QY 1540 TCTTGGAGATGCTAATCTCT 1562
DB 1518 TCTCTGAGATGAAGATCTTT 1540

RESULT 13
LOCUS PSU34744 1799 bp mRNA linear PLN 02-DEC-1996
DEFINITION Phalaenopsis sp. 'hybrid SW9108' cytochrome P-450 (CYP8A2) mRNA,
complete cds.
ACCESSION U34744
VERSION U34744.1 GI:1173623
KEYWORDS
SOURCE
ORGANISM Phalaenopsis sp. SW9108
Phalaenopsis sp. SW9108
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Asparagales; Orchidaceae;
Epidendroideae; higher Epidendroideae; Vandeeae; Aeridiinae;
Phalaenopsis.
REFERENCE 1 (bases 1 to 1799)
AUTHORS Nadeau,J.A., Zhang,X.S., Li,J. and O'Neill,S.D.
TITLE Ovine development: identification of stage-specific and
tissue-specific cDNAs
JOURNAL Plant Cell 8 (2), 213-239 (1996)
MEDLINE 96351416
PUBMED 8742709
REFERENCE 2 (bases 1 to 1799)
AUTHORS Nadeau,J.A.
TITLE Direct Submission
JOURNAL Submitted (23-AUG-1995) Jeanette A. Nadeau, Section of Plant
Biology, University of California, Davis, Davis, CA 95616, USA
FEATURES
source 1..1799

```

/organism="Phalaenopsis sp. SM9108"
/mol_type="mRNA"
/strain="hybrid SM9108"
/db_xref="taxon:54872"
/clone="O40"
/tissue_type="pollen tubes"
/clone_lib="megaportocyte-stage ovule lambda Zap11
library"
1.1799
/gene="CYP78A2"
291.1571
/gene="CYP78A2"
/codon_start=1
/product="cytochrome P-450"
/protein_id="AA87221.1"
/db_xref="GI:1173624"
/translation="MAFSVGLTRFVSSHRTAKELSPAPADRPKESAVELLFNR
AMGAPAPGDYWRNLRISSTYLSPREVSEFEXORSEIGEMTRDMKMERGVVEV
RMRLVYSLNNIMLVGKKEDPAKDELESLIKGYELGIFMNGHLPLGMLD
LOGVRRCTLVAKNVFVKKIIDEHRRANGVIGDEGEDPVDVLGLEEDRLSE
SDMVAVMEMIPEGTDTVALILEMTLAPMLADPDLOSROVEIDVDSRPLDSDI
ORLPYLOSIVKETRMRHPPGLPISMARLAIHDVPDGHNPACTTMMVNMATTHDEC
NMAEPKFNDFRFLDEVDNLISDLRLAPGSGKRVCPGKTMLAAVHMLQDLKSF
KLPSSRNVLSBCLKMSLEKNPLVCVAVRPE"
BASE COUNT 456 a 364 c 428 g 551 t
ORIGIN

Query Match 22.3%; Score 355; DB 8; Length 1799;
Best Local Similarity 58.0%; Pred. No. 2.8e-96;
Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTCGAGCCCAAGAGCTCATGCTTTTCAGCTTAGAGAGAGCTGCGGTGATGTCAGGTGC 363
DB 273 TTGAGGCCCTCGGCTCATGCTTTCTCGGTGGGCTCACTGCTTCATCGTTTGAAGC 332
QY 364 AATCCGACGTAGCGAAGAGATTCGAATAGCCCGGTTTTTGTCTATGCAACCGGTTAA 423
DB 333 CACCCGAAACCGGAAAGAGATTCCTTCAAGCCACCTTCGTATCGGCCCATTTAA 392
QY 424 GAATCGCTTACTCACTGATGTTTAAAGAGCAATGTTTGAACACACGCGTTTAC 483
DB 393 GAATCAGATACGAACCTCTGTTTAAATCGGCTATGGGTTTTGCCCATTTGGGATTC 452
QY 484 TGGCGAACGCTTCCCGCTATCGCTTGAACCATCTCTTTAGTACAAACAAATCAGAGA 543
DB 453 TGGAGAAACCTGAGAGAGATTCGTCCACATATCTTTTCAGTCCGCGCAGATTCATCG 512
QY 544 GCCGAGACGACAGACAGATGATCTCAAGCAGATGGTGTGATTTCTTGAAGAAACAGA-- 601
DB 513 TTGCGAAGACGACGAGTGAATGGCGAAGAGATGTTGGGAGTATGAAAGAAATGATG 572
QY 602 -GTAGTAACGAAACCTGTTTTTGTTCGTGAGTTCCTTAAACGCGCTGCTTAAACAATG 660
DB 573 GAGAGAAATGAGATTGTAGAGTGAAGAGATGTGCACTACGCGCTCTTTGAATTAACATC 632
QY 661 ATGTGCTCTGATTCGACAAAGATGATGAGTCTTAAAGAAACCACTGTGAGTAACTGAA 720
DB 633 ATGTGACTGTTTTTGGAAAGAGTGTGATTTTCCAAAGGATGAGGGGTTGAGACTTGAG 682
QY 721 ATGTGC--GAAGAGGTTATGATTTGCTCGGAAAGTGAATTGAGTATCACTTCT 777
DB 693 TTGATCCTTAAGAGAGTATGAGTACTTGAGATCTTCAACTGGGATGATCACTTCCCT 752
QY 778 TGGCTATCGAGTTGATCTCTCAAGACTCCGCTGATGATTTCCAAATCCTACCAAG 837
DB 753 CTTTGGAGATGATTTGCAAGGATGAGTGAAGAGAGTACAGAAACCTTGTGGCTAAG 812
QY 838 GTAAACCGGTTTGTATCCCGGATTAATCCGAAACCGTATCAACCC----- 885
DB 813 GTCAATGATTTTGAAGAGATCATAGACAGATTAAGAGAGCAACCGGCTAGGG 872
QY 886 -----GGTGAATGCTCTGATTTGCTCAAGCTTTTGTCTCTCCCTCCATGATTCAT 939

```

```

DB 873 ATTGATGAGGTGAAGGTGAAGATTTTGTGATGTCCTTCTGGTTGGAGAGAAAGAT 932
QY 940 AAATTAATCCGACCCCGACATAATGCGCTTTTGGAGATGATATTCAAGAGAACAGAC 999
DB 933 AGACTCTGAGATCTGATATGATGTCGCAAGTTCTTTGGGAAAGATCTTTAGAGAACTGAT 992
QY 1000 ACAGTTGGGCTTAAATGAGAGATCTCGCTGAGATGTCCTTCATCCAGATATGCAA 1059
DB 993 ACTGTGCGATCTTATGGAATGAGCTGGCTAGATGATTTCTTCATCTCGATTTCAA 1052
QY 1060 TCACGCTACAAACAGCTGATCACTAGTATGAGGAAATCAAGACCTTAGATGATCT 1119
DB 1053 TCGAAGGACAAAGTGAATGATTTCTGTGCTGCTTCAAGGCCAGTATGGAATCT 1112
QY 1120 GACTTGGCTTCACTTCCATCTATCAAGCGCTGAGTGAAGAGATTTAGGCTTCATCT 1179
DB 1113 GATATCCAAAGCTTCTTATCTCCATCTATAGTAAAGAAACCTTCGATGATCT 1172
QY 1180 CCAGGCCCACTTCTATCATGAGGCCCGTTGGCCATTAACAGACAGATGTTGATGCTGT 1239
DB 1173 CCGGGGCTCTATTTGTCATGGGCTCGCTAGCTATCCATGACGTTCTGTTGATGCTAC 1232
QY 1240 CTTGTTCCGGCAGGAGCAACAGCAATGATGAACATGTGGCCGTTATGCAATGATCCACAC 1299
DB 1233 ATGATTCCTGTGAGGACAGCTGCAATGATGTAACATGTGGCAATTAACACATGACGAATGC 1292
QY 1300 GTGTGGGTGATCTTGTGAAGTTTAACTGAGAGTTCGTGGCAAAAGAGTGAGGTG 1359
DB 1293 AACTGGCTGAGCTTAACAAATTCATCTGATGATTCATCGATGGAAGATG----- 1344
QY 1360 GAGTTTCGCTTCTGAGTTCGAGTTGAGACTTGACCTTTCCGGTCCGCTCGATTT 1419
DB 1345 ----TCAATATCTTGTTCGATTTAAGTGTGACCTTTGGCTCCGGAAGAGATT 1400
QY 1420 TGGCCGGGAGAAATCTTGTGTTTACTACCGTTATGTTTGAAGGAGATGATTTACAT 1479
DB 1401 TCCCTGGCAACATGATGATGATGCTGCAATTCATCTTGTGCTCACTTCTGAAA 1460
QY 1480 GAGTTTGAATGGGACCGTCCGATGTAAGCGCTTGAATCTATGAGAACTAGAGCTT 1539
DB 1461 AGCTTCAATGCTTCTTCTC--GAGAATGATGATGATTTGCTGAGTGCCTAAAGATG 1517
QY 1540 TCTTGGAGATGCTAATCTCT 1562
DB 1518 TCTCTCGAGATGAGAAATCTTT 1540

```

```

RESULT 14
AF049067 1990 bp mRNA linear PLN 06-WAR-1998
LOCUS Pinus radiata cytochrome P450 (PHE74) mRNA, complete cds.
DEFINITION
ACCESSION AF049067
VERSION AF049067.1 GI:2935524
KEYWORDS
SOURCE
ORGANISM
Pinus radiata (Monterey pine)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Coniferopsida; Coniferales; Pinaceae; Pinus; Pinus.
1 (bases 1 to 1990)
REFERENCE
1 Bishop-Hurley,S.L., Walter,C. and Gardner,R.C.
Isolation and Expression of Abundant mRNAs during Somatic
Embryogenesis of Pinus radiata
Unpublished
2 (bases 1 to 1990)
REFERENCE
Bishop-Hurley,S.L., Walter,C. and Gardner,R.C.
Direct Submission
Submitted (18-FEB-1998) Biotechnology Division, NZ Forest Research
Institute, Sala Street, Rotorua, New Zealand
1.1990
Location/Qualifiers
FEATURES
source
/organism="Pinus radiata"
/mol_type="mRNA"
/db_xref="taxon:3347"

```

```

/tissue_type="somatic embryonic"
/dev_stage="stage 1 embryo"
1..1990
/gene="PRR74"
3..1734
/gene="PRR74"
/codon_start=1
/product="cytochrome P450"
/protein_id="AAC05148.1"
/db_xref="GI:2935525"
/translation="MENRBSGGSGMMVVCVLPFTKDPATFLHSSSDVSNRQWPL
YIALIVHGAQKQMSFSLGCTPAVITSDPVARBLSPHPANPLKQSAQOLFGRA
LESLARLHGAQKQMSFSLGCTPAVITSDPVARBLSPHPANPLKQSAQOLFGRA
IGFAPNGSTWRLRLRIASAHLPAPRIIAHAGQAVVAMLDIOEYHRSKVPR
RHLQALANIMSVFGRFDMSHENREKREMBVDFGLIAFWADHLPWLRPL
DPLRIHARCARLVPRTVFSNII EQRREORESDQDFVVLISLQGEKDEE
DMIAVLEMIERGLDTALITMTWALVLEHPEAKQAKQELDVGHDSVXDSDIP
KLPIQAVVKEALRKHPPGLISWALSTEDVNMGDGCVAPAGTAAVNMMSITHDN
IMESPYERPRFVFEVGEEDVGRANDRLAPGAGRVCPCGALGLATVNLVATL
LHFEWLPFAHEPDLSEVLTSCENARPLKCVPTVTPFAKESD"
BASE COUNT      408 a      472 c      560 g      550 t
ORIGIN
Query Match      21.9%; Score 348.2; DB 8; Length 1990;
Best Local Similarity 55.5%; Pred. No. 3.3e-94;
Matches 768; Conservative 0; Mismatches 588; Indels 27; Gaps 4;
211 ATCCCGGTCCAAAGGCTTCCCTTGGTGAAGCATGTCACTCATGTCAAGACCTTA 270
307 ATTCCTGGCCCCAAAGATGGCCGATTTATGGAACTTAATGACATAGAGCGTGGCTT 366
271 GCTACCGACGATCGCTGATGACGTGAGAAATTCGAGGCCAAGAGCGCTATGGCTTC 330
367 CCTCACCGAAGCTTGAATCTCTCGCTCTTCAATGATGGAGAGCAAGCTCATGTCCCTT 426
331 AGCTTAGAGAGACTCGCGTATCGTACGTGCAATCCCGACGTAGCGAAGAAGATTCTG 390
427 ACCTGGGCTGACCGCCCGCTGTATTAAGTCGATCCCGAGGTGGCCGAGATGGCTC 486
391 AATAGCCGGTTTTGCTGATGACGCGTTAAAGATCGGCTTACTCATGATGTTTAAAC 450
487 ACCTCCCTCACTTGCACACAGACCCCTCAAGCAATCGCACAGACAGTCTGTTGGG 546
451 AAGACAAATGTTTGGACACACGCGTGTTACTGCGCAACGCTTCCCGATATGCTTGG 510
547 AAGGCTATTGGCTTGCACCCCAACGCGGCTACTGAGGCTGCTCAGAGAAATTCCTCT 606
511 AACCATCTTTAGTCAAAACAATCAAGAAGCCGAGACGCAACGACGATATCTCA 570
607 GCTCATCTCTTGGCTCTCGACAGATGCTGCCCATGAAGCTGGCCGCGACGACCTT 666
571 AGCCAGATGTTGAGTTTCTTGA AAAACAGAGTATGACAAACCTGTTTT--GTTCGT 627
667 GTGGCAATCTGATGATATCCAAAGAGTATCAATTAAGGTTGTTTATGAGTTAAGA 726
628 GAGTTCCTTAAACGCGCTCGCTTACACATGATGTCTCTGATATGCGAACAAGATAT 687
727 CGACACCTGCAGACCGCGCTGAAACAATATCATATGGAAGGCTCTTGGAGAGGTTTC 786
688 GAGCTTGA AAAAACCATGTTG-----AGTTACGTGAATATGATCGAAGAGTTAT 738
787 GACATGTCTCATAGATAGAAAGAAAGTGAAGAGCTCGGAGAGATGTTGATGAAGGTTTC 846
739 GATTGCTCGGAAGCTTGAATGAGTATCACTTCTTGGCTATCGAAGTTTATCTCT 798
847 CAAGTGTGAGACGCTTCACTGCGGAGATCATCTCCATGCTGCGACCGCTGACCTT 906
799 CAAGAAGCTCGGCTCATAGATTTCCACACTGTATCAACAAAGTAAACCGGTTTATCCCG 858
907 CTTCGATTCATGCTCGCTCGCTCGCTCGCTCGTACCGCGTGTGCAACTTTCGTTAGCAT 966
859 ATTATATCCGAACCGCTA-----ATCAACCGGCTGATTTGCTCTGTATTTTC 906

```

```

Db 967 ATCATCGAAGACACCGCTCGGGAAGAGCGGTAGAGAAAGCGGATCAATGTGATTT 1026
Qy 907 GTGACGTTTTCCTCTCCCTCCATGTTCAATTAATATCGACCCGACATATATGCC 966
Db 1027 GTTGAAGTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
Qy 967 GTTCTTTGGAGATGATATATTCAGAGGAACAAGACAGTGGGCTTATGAGTGAATC 1026
Db 1087 GTTCTCTGGAATATATCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1146
Qy 1027 CTGCTGAGATGCTCTCTTATCCATATGCAATACGATACAGTACAAAGAGCTGATCAA 1086
Db 1147 ATGGCAGAACTGTTCTGACCTTGAAGCTGAGAAAGAGTCAAGAGAGTGAAGCC 1206
Qy 1087 GTAGTGGGACAGACCGAAGCTGAAGATTCAGACATTCAGAAAGCTGCCATACATCAA 1266
Db 1267 GAGTGTGAAGAGCGCTGCGAATGATCCACCTGACCTCTTATCATGGGCTGCT 1326
Qy 1207 TTGGCCATTAACA---GACATCGTTGATGCTGCTTCTGCTGCTGCTGCTGCTGCTGCT 1263
Db 1327 CTCTTACAGAAAGATGTAACATGGGCGATGAATGTGTGTTTCAAGGCGACAACTGCC 1386
Qy 1264 ATGGTAACATGAGGCGCGATATGCAATGCAACAGTGTGGATGATCTTGAAGTTT 1323
Db 1387 ATGGTAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1446
Qy 1447 CGTCCGAGAGGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1506
Db 1384 TTGAGACTTGCACCTTTCGGGCTCGGCTCGGATTTGCTCCGAGAAATCTGATTTT 1443
Qy 1507 CTGAGCTTCCGCTTCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1566
Db 1444 ACTACGTTATGTTTGAACGCGGATGATGATGATGATGATGATGATGATGATGATGAT 1503
Qy 1567 GCCACTGTTATCTCGGTTGGAAGTGTCTCCATTTGAATGAGTCTTCCATATGCT 1626
Db 1504 GGTAGCGGTTGATCTGAGAACTGAGGCTTCTTGGAGATGCTATCTCTT 1563
Qy 1627 GAACATCGGTTATCTGCGAGTCTTGAAGTCTTGAAGTCTTGAAGTCTTGAAGTCTT 1686
Db 1564 CCT 1566
Db 1687 CAT 1689

```

RESULT 15
 AL772960/c 565 bp DNA 1linear STS 11-JUN-2003
 LOCUS Arabidopsis thaliana transposon insertion STS SM_3_556, sequence
 DEFINITION tagged site.
 ACCESSION AL772960
 VERSION AL772960.1 GI:21535147
 KEYWORDS STS; STS; sequence tagged site.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosoids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS Clarke, J.H., Bowles, B., Carter, J., Hart, D., McCullagh, B.,
 Murphy, G., Jangnam, S., Leary, C., Jones, J.D.G. and Bevan, M.
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 565)
 AUTHORS Clarke, J.H.
 DIRECT SUBMISSION Submitted (18-JUN-2002) Clarke J.H., John Innes Centre, Colney
 Lane, Norwich, NR4 7UJ, UK

COMMENT

AT denotes an activation tag dissociation transposon within a single line, ET an enhancer trap dissociation transposon, GT a gene trap dissociation transposon, MT a mis-expression enhancer trap dissociation transposon, SW a defective suppressor mutator transposon. 3 denotes a sequence derived from the 3' end of the transposon, 5 denotes a sequence derived from the 5' end of the transposon. BBSRC GARNET, ATIS project
On-line seed stock requests: <http://naec.mot.ac.uk/NASC/stock>
code: N5511.

FEATURES

source

Location/Qualifiers
1..565
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/variety="Columbia-0 NASC stock code N1092"
/db_xref="taxon:3702"
/clone="AL138642"
/note="Derived from superpool 12.34 NASC code Unknown"
1..565
/standard_name="SM_3.556"

STS

BASE COUNT 167 a 151 c 112 g 135 t
ORIGIN

Query Match 21.1%; Score 336.2; DB 11; Length 565;
Best Local Similarity 83.9%; Pct: No. 1.3e-90;
Matches 380; Conservative 0; Mismatches 73; Indels 0; Gaps 0;

975 GGAGATGATATTCAGAGGACAGACAGATGCGGCTTAAATGAGTGAATCTCGCTAG 1034
453 GGAAATGATATTCAGAGGACAGACAGATGCGGCTTAAATGAGTGAATCTCGCTAG 394
QY 1035 GATGTCCTTATCCAGATATGATCAACGCTACAAACGAGCTGATCAAGATCGG 1094
DB 393 GATGTCCTTATCCAGATATGATCAACGCTACAAACGAGCTGATCAAGATCGG 334
QY 1095 GAATCAAGAGCCCTAGATGATCTGACTTGCTTCACTTCATATCTAAAGGCTGTGT 1154
DB 333 ACGATCAAGGCTGTGAGAGGTCTGAGGTGTCTAGTATATCTAAAGGCTGTGT 274
QY 1155 GAAGAAGATGAGGCTTCACTCCAGGCCCACTTATCAATGAGGCCGCTTGCCAT 1214
DB 273 GAAGAAGCTTGAAGCTTCACTCCAGGCCCACTTATCAATGAGGCCGCTTGCCAT 214
QY 1215 AACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
DB 213 CAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 154
QY 1275 GTGGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
DB 153 GTGGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 94
QY 1335 GTTCGTGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1394
DB 93 TTTGTAGCAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 34
QY 1395 ACCCTTCGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1427
DB 33 ACCGTTGGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGG 1

Search completed: January 17, 2004, 03:16:35
Job time : 5890 secs

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OW nucleic - nucleic search, using sw model

Run on: January 16, 2004, 09:48:47 ; Search time 124 Seconds
(without alignments)

7151.119 Million cell updates/sec

Title: US-10-022-025a-1

Perfect score: 2009

Sequence: 1 aataaataaataaatacgtta.....atagagattagagatttt 2009

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/ptodata/1/ina/5A COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	420.8	20.9	1611	3	Sequence 7, Appl
2	200.4	10.0	1799	2	Sequence 3, Appl
3	123.1	6.1	1704	3	Sequence 17, Appl
4	101.2	5.0	1708	3	Sequence 1, Appl
5	98.2	4.9	275	4	Sequence 2591, Ap
6	89	4.4	1788	3	Sequence 9, Appl
7	88.8	4.4	1893	1	Sequence 1, Appl
8	84.6	4.2	2174	3	Sequence 63, Appl
9	84.6	4.2	2174	3	Sequence 63, Appl
10	83.2	4.1	1761	3	Sequence 3, Appl
11	83.2	4.1	1927	3	Sequence 64, Appl
12	83.2	4.1	1927	3	Sequence 64, Appl
13	81.4	4.1	5156	2	Sequence 3, Appl
14	81.4	4.1	5156	2	Sequence 3, Appl
15	79.2	3.9	1722	3	Sequence 1, Appl
16	77.2	3.8	1755	1	Sequence 29, Appl
17	77.2	3.8	1755	1	Sequence 29, Appl
18	77.2	3.8	1755	1	Sequence 29, Appl
19	76.4	3.8	1812	1	Sequence 37, Appl
20	76.4	3.8	1824	3	Sequence 1, Appl
21	76.4	3.8	1824	3	Sequence 1, Appl
22	75.6	3.8	384	4	Sequence 232, App
23	75.6	3.7	1835	4	Sequence 5, Appl
24	75	3.7	1838	4	Sequence 1, Appl
25	73	3.6	1620	4	Sequence 11, Appl
26	73	3.6	1620	4	Sequence 11, Appl
27	73	3.6	1620	4	Sequence 11, Appl

28	72.8	3.6	543	4	US-09-615-192A-230	Sequence 230, App
29	71.8	3.6	1894	4	US-09-564-808-3	Sequence 3, Appl
30	70.2	3.5	1657	3	US-08-948-564-11	Sequence 11, Appl
31	69.2	3.4	1806	4	US-09-351-229-3	Sequence 3, Appl
32	68.6	3.4	1880	4	US-09-564-808-1	Sequence 1, Appl
33	66.6	3.3	1781	4	US-09-499-302A-1	Sequence 1, Appl
34	65.4	3.3	1812	1	US-07-912-900-28	Sequence 28, Appl
35	65.4	3.3	1812	1	US-08-285-309-28	Sequence 28, Appl
36	65.4	3.3	1812	2	US-08-502-046-28	Sequence 28, Appl
37	64	3.2	1506	3	US-09-158-767-7	Sequence 7, Appl
38	64	3.2	1506	3	US-09-158-767-8	Sequence 8, Appl
39	64	3.2	1506	3	US-09-158-767-9	Sequence 9, Appl
40	64	3.2	2261	3	US-09-158-767-1	Sequence 1, Appl
41	63	3.1	1737	4	US-09-126-420A-1	Sequence 1, Appl
42	63	3.1	1824	3	US-08-948-564-13	Sequence 13, Appl
43	62	3.1	386	1	US-08-313-075A-55	Sequence 55, Appl
44	61.4	3.1	1634	4	US-09-126-420A-2	Sequence 2, Appl
45	61.4	3.1	1644	3	US-08-948-564-5	Sequence 5, Appl

ALIGNMENTS

RESULT 1
US-08-948-564-7
Sequence 7, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
TITLE OF INVENTION: Cordin, Frederick T.
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESS: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..1588
US-08-948-564-7
Query Match 20.9%; Score 420.8; DB 3; Length 1611;
Best Local Similarity 58.1%; Pred. No. 4.9e-107;
Matches 1029; Conservative 0; Mismatches 487; Indels 256; Gaps 6;

221 AGCCAAACCAACCTTGGCTTCTCCCTCTGCGCGTCAATCATCTGGCTGGCATATCT 280
74 ACCCAAGAAAACCTTGATGAGGTCTTTTATGATGAGGCTCACTCTGTAAACATGACT 133
281 CTCTCTTATGAGACTTATCCCCGTGACCTGCTGGGGGAAATACCTCTTGGCCGGTAA 340
134 TTCTATTACTGTGACACCCCGGTGCTGCTGGGGCAAGTAACTACAC-----TA 186
341 ATATCCGGTTCATACAAAACCGGAAAGTATTTCCCGGTCCAAAAGGCTTCCCTTGGTT 400
187 CTCTCCCCCTTTCAA-----TCAATCCCGGTCCAAAAGGCTTCCCTTAAAT 235
401 GGAAGCATGTCACTCATGTCAAGCACTAGTCAACGAAATGCTGTATGAGCTGAG 460
236 GGAAGCATGGGCTCATGACTTCC--CTGGCCCATACCGTATCGACGCCGGCCGCC 292
461 AAATTCGAGCCAAAGGCTCATGCTTTCAGCTTAGAGAGACTGCGTGAATGTCACG 520
293 ACATGCAAGCCCAAGCCCTCATGCTTATGCTGGGCAACAGGTGTCATGTCACG 352
521 TGCAATCCGACGTAGCGGAAAGATTTCTGAATGCCCCGTTTTTGTCTGATCGACGG 580
353 TGCCACCCCGACGTGGCCCAAGAGATTTCTCAACAGCTCCGTCTTCCCGCATGCTCC 412
581 AAAGATCGGCTTCTCATCATGATGTTTAAACAGACCAATTTGGTTCACACACGGTGT 640
413 AAAGATCGCATACAGCTCATGTTTAAACGCGCCATCGGCTTCCGCTTTAGAGATT 472
641 TACTGGGAAAGCTTCCGCGTATGCTTCCGAAACCATTTCTTATGATCAAAACATGAG 700
473 TACTGGGAAAGCTTCCGCGTATGCTTCCGAAACCATTTCTTATGATCAAAACATGAG 532
701 AGAAGCCGAGGCAAGCAAGCATGATCTCAAGCAGATGCTGATTTCTTGAAGAAACAG 760
533 GCTCTGAGCTCCAAAGCTCTCAAAATGCGGCCCAAAATGTTTCAATCTCT---AAATAC 589
761 AGATGTAACGAACCTCTTTTGTTCGTGAGTGGTCTTAAACGGCGTCTTAAACATG 820
590 AAGCGCCACCGCATGCTTACGTGTGCGCAAGTCTGTAAGAAAGGCTTGTCTCATGATG 649
821 ATGCTCTGATATTCGGAACAAGATGATGAGTGAAGAAACCA-----TGTGAG 871
650 ATGCTCTGATATTCGGAACAAGATGATGAGTGAAGAAACCA-----TGTGAG 871
872 TTAAGTGAATGCTGGAAGAGTATGATTTGCTCGAAGCGTGAATTTGAGCTGATCAC 931
710 CTGGAATATTAGTGAACCAAGTATGACCTGTGAGGCTGTTTAATGAGCGGACAC 769
932 CTTCCTTGGCTATGAGAGTGTGATCTCAAAAGCTCGGCTTAATGTTCCACCTGCTA 991
770 CTTCCTTGGCTATGAGAGTGTGATCTCAAAAGCTCGGCTTAATGTTCCACCTGCTA 829
992 CCAAGGTAACCGGTTTGTATCCCGGATTAATCCGAAACCGTAAATCAACCGGTGAT 1051
830 CCGATGATGAGACCGTTTGTGCGGACATCATGCTGAACACCGAGCTAGTAAACCGAA 889
1052 TTGCTCTGATTTGCTGAGCGTTTGTCTTCCCTCCATGCTGATGATTAATTTCCGAC 1111
890 ACCAATCGGATTTTGTGAGCTCTTGTCTCTCCCGGAACTGATCAATTTATGAGC 949
1112 CCGGACATATGCGGCTCTTGTGGTATGACACACCATTAATTTGATTAATTTCTTAAT 1171
950 TCCGACATGATGCTGATCTTG-----972
1172 TATATTTTGTGAAAAATGCTTAGATTAATTAGTAAACATGAAATTTGAGCTCAAT 1231
973 -----972
1232 GTGAGGTGTTGGAATAATTAAGCATTAAGAAATTTTGTGATCAATTAATTAATTA 1291
973 -----972
1292 ATTTAGATTTTAAACAGTTTCTATATAAAAGTAAAGTAAATTTTGTGATTAATTT 1351

973 -----972
1352 ATTAAGAGATGATTAATCAAGAGACAGACAGATTTGGCTTAAATCAAGTGAATCT 1411
973 -----GGAATGATTAATCAAGAGACAGACAGATTTGGCTTAAATCAAGTGAATCT 1026
1412 CGTGAATGATGCTTCAATCAAGATTAATCAAGATTAATCAAGATTAATCAAGATTA 1471
1027 CGGAGATGAGCGCTTCAATCTCAATGAGTCAATCAAGATTAATCAAGATTAATCA 1086
1472 AGTCGGAATTAAGAGCGCTTGAATGAATGATGATGATGATGATGATGATGATGAT 1531
1087 TGTGGAAGACAGCGCGCTCGCAGAGATGATGATGATGATGATGATGATGATGAT 1146
1532 TGTGGAAGACAGCGCGCTCGCAGAGATGATGATGATGATGATGATGATGATGAT 1591
1147 GGTGGAAGAGAGGCTGCGCTGACACCGCGCGCTCACTTATCAATGAGCTGCT 1206
1592 GGCATTAACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1651
1207 GTCCATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1286
1652 GAACATGAGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1711
1267 CAACAGTGGGCTATTTGCAAGGACCCACAGTGTGAAGAGACCATCGAATTTATG 1326
1712 TGAGAGTTCGTCGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1771
1327 CGAGAGTTCGTCGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1386
1772 ACTTCACCTTTGGGCTGCGCTGCTGATTTGCGGCTTCCGCTTGGCTGATTTGAG 1831
1387 ACTTCACCTTTGGGCTGCGCTGCTGATTTGCGGCTTCCGCTTGGCTGATTTGAG 1446
1832 CGTATGTTTGAAGAGGAGTGAATGATGATGATGATGATGATGATGATGATGATGAT 1891
1447 GGTGAATTTTGGGCTGCGCTGCTGATTTGCGGCTTCCGCTTGGCTGATTTGAG 1506
1892 CGGCTGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1951
1507 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1566
1952 TAAATGCGCGCTGAGCGGCTTAAATTAATTAATTAATTAATTAATTAATTAATTA 1983
1567 CAAAGTGGCGCGGCTGATTAAGAGAG 1598

RESULT 2
US-08-560-398-3
Sequence 3, Application US/08560398
Patent No. 5907082
GENERAL INFORMATION:
APPLICANT: O'Neil, Sharmen
APPLICANT: Nadeau, Jeanette
TITLE OF INVENTION: Ovine-Specific Gene Expression
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/560,398
FILING DATE: 17-NOV-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Bastian, Kevin L.
REGISTRATION NUMBER: 34,774
REFERENCE/DOCKET NUMBER: 023070-063300US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1799 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: unsure
LOCATION: 19..23
FEATURE:
NAME/KEY: CDS
LOCATION: 291..1571
OTHER INFORMATION: /note= "clone O40 ovule-specific gene
OTHER INFORMATION: encoding a cytochrome P450 monooxygenase
OTHER INFORMATION: from pollen tubes of Phalaenopsis"
US-08-560-398-3

Query Match 10.0%; Score 200.4; DB 2; Length 1799;
Best Local Similarity 61.6%; Pred. No. 9e-46;
Matches 362; Conservative 0; Mismatches 211; Indels 15; Gaps 2;

QY 1358 GGAGATGATATTCAGAGGAGACAGACAGTGGCGTCTTATGAGTGAATCCTCGCTAG 1417
DB 968 GGAATATATTTAGAGAGACTATCTGTGGCACTCTTATGAGTGAATGAGCTGGCTAG 1027
QY 1418 GATGCTCTTCATCCAGATATGCAATCAACGATACAAAAGAGCTGATCAAGTATCGG 1477
DB 1028 AATGCTCTTCATCCAGATATGCAATCAAGAGCAAAAGTGAATGATCTGTCTGTA 1087
QY 1478 GAATCAAGAGCCCTAGATGATCTGCTTCACTTCCATCAATCAAGCTGTGTGT 1537
DB 1088 CTCTCAAGCCAGATGATGATCTGATCTCAACGATCTCTTATCTCAATCTATAGT 1147
QY 1538 GAAAGAGATATGAGGCTTCACTCTCCAGGCCCACTTCAATCAATGAGCCGTTGGCAT 1597
DB 1148 AAAAGAAACCTTCGATGATGATCTCTCGGCGCTTATGATCAATGAGCTGCTAGCAT 1207
QY 1598 AACAAGACAGATGATGATGATGATCTGTCTTCCGAGAGGACCAAGCAATGATGAAT 1657
DB 1208 CCAATGAGGTTCTGTTGATGATGATCAATGATCTCTGCTGAGCAAGTGAATGATGAAT 1267
QY 1658 GTGGGCGGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1717
DB 1268 GTGGGCGGATTAACATTAACATTAACATTAACATTAACATTAACATTAACATTAACAT 1327
QY 1718 GTTCTGCGCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777
DB 1328 ATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1375
QY 1778 ACCCTTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1837
DB 1376 ACCCTTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1435
QY 1838 GTTGGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1897
DB 1436 TCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1492
QY 1898 TGACTTATGAGAACTGAGGCTTCTTGGAGAGATGATGATGATGATGATGATGATGATGAT 1945
DB 1493 AGATTTGCTGAGTCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1540

RESULT 3
US-08-948-564-17
Sequence 17, Application US/08948564

Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Simonsky, Balazs
APPLICANT: Dewey, Ralph E.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 38..1564
US-08-948-564-17

Query Match 6.1%; Score 123.2; DB 3; Length 1704;
Best Local Similarity 54.3%; Pred. No. 2.5e-24;
Matches 301; Conservative 0; Mismatches 228; Indels 15; Gaps 2;

QY 1358 GGAGATGATATTCAGAGGAGACAGACAGTGGCGTCTTATGAGTGAATCCTCGCTAG 1417
DB 922 GGAATATATTTAGAGAGACTATCTGTGGCACTCTTATGAGTGAATGAGCTGGCTAG 981
QY 1418 GATGCTCTTCATCCAGATATGCAATCAACGATACAAAAGAGCTGATCAAGTATCGG 1477
DB 982 GTTCTGCGCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1041
QY 1478 GAATCAAGAGCCCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537
DB 1042 GTTGAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1101
QY 1538 GAAAGAGATATGAGGCTTCACTCTCCAGGCCCACTTCAATCAATGAGGCGGTTGGCAT 1597
DB 1102 CAAGAGCAATGAGGCTTCACTCCAGGCCCACTTCAATGAGGCTTCAATGAGGCTTCAATGAG 1158
QY 1538 AACAAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1657
DB 1159 TGCCATGATCAAAAGTTGAGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218
QY 1658 GTGGGCGGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1717
DB 1219 GTGGGCGGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1278
QY 1718 GTTCTGCGCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1777

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rni

Page 4

Db 1279 GTTCCTTGAGAGATGTAGCATGAAG-----GGCCATGACTTTAGGCTACT 1326
Qy 1778 ACCTTCGGGTGGGTGCTGCGATTTGCCCCGGGAGAAATCTTGTTTACTACCTGTTAT 1837
Db 1327 TCATTGCGGTGGGTGAGAGATGCGCGGTGGCCAACTTGATCAACTGGCAGC 1386
Qy 1838 GTTTGGACGGCCATGATGTACATGATTTGAATGGGACCGTCCGATGATGAGCGGT 1897
Db 1387 ATCCATGTTGGGCGCACCTCTTGACCACTTCTGTGGACCCCACTGAAGGATGAAGCC 1446
Qy 1898 TGACTTATCTGAGA 1911
Db 1447 TGAAGAAATGACA 1460

RESULT 4
US-08-991-677-1
Sequence 1, Application US/08991677A
Patent No. 6252135
GENERAL INFORMATION:
APPLICANT: Chiang, Vincent L
APPLICANT: Carriaway, Daniel T
APPLICANT: Smeltzer, Richard H
TITLE OF INVENTION: Production of Synglyl Lignin in Gymnosperms
FILE REFERENCE: 50617
CURRENT APPLICATION NUMBER: US/08/991,677A
EARLIER FILING DATE: 1997-12-16
EARLIER APPLICATION NUMBER: US 60/033,381
NUMBER OF SEQ ID NOS: 11
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 1
LENGTH: 1708
TYPE: DNA
ORGANISM: Liquidambar styraciflua
FEATURE:
NAME/KEY: CDS
LOCATION: (48) .. (1571)
US-08-991-677-1

Query Match 5.0%; Score 101.2; DB 3; Length 1708;
Best Local Similarity 52.4%; Pred. No. 3,2e-18;
Matches 279; Conservative 0; Mismatches 238; Indels 15; Gaps 2;

Qy 1358 GAGATGATATTCAGAGAGACAGACAGATTCGAGTTCATTCAGATGATTCCTGCTAG 1417
Db 929 GATATGATCATCTGAGGATGAGACACACACGCAATCTCTGCAATGGGCGCATGGCGCA 988
Qy 1418 GATGCTCTTCATTCAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1477
Db 989 GTTAATTAAGAACCCAGAGGTGACACAAAGCTTCAGAGAGATGATGATGATGATGATGAT 1048
Qy 1478 GAAATCAAGAGCCCTAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1537
Db 1049 GTCCGAGCGTGTCTGACCGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1108
Qy 1538 GAAAGATATTCAGAGGCTTCATTCCTCAGGCGCCATTCATTCATTCATTCATTCATTCAT 1597
Db 1109 CAGAGAGGACTTAAGGCTGACCCCTTCA--ACACCATTAATGCTCCTCTCATCGCGGCA 1165
Qy 1598 AACAGACAGATCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1657
Db 1166 TGCCAAAGCTCAAAATTTGGGTGCTAGACATCCCTTAAGGATCAATGTTGATGATGATGAT 1225
Qy 1658 GTGGGCGGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1717
Db 1226 CTGGGCGGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1285
Qy 1778 GTTCGTGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1777
Db 1286 GTTCTGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1333
Qy 1778 ACCTTTGGGTGGGTGCTGCGATTTGCCCGGGAAGATCTGTGTTTACTACCGTTAT 1837

Db 1334 GCGGTGTTGTCAGAGAGCGGTGTTTCCCGGTGACAACTTGACATTCATTTGGTTCAC 1393
Qy 1838 GTTTGACGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1889
Db 1394 ATCCATGATGGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445

RESULT 5
US-09-313-294A-2591
Sequence 2591, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 2591
LENGTH: 275
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Inocyte ID No. 6476212 700552945H1
US-09-313-294A-2591

Query Match 4.9%; Score 98.2; DB 4; Length 275;
Best Local Similarity 64.7%; Pred. No. 9.1e-18;
Matches 178; Conservative 0; Mismatches 93; Indels 4; Gaps 2;

Qy 1582 GGGCGCGTGGCCATTAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1641
Db 5 GGGCGCGCTCCACCTCAACAGCTGACAGCTGATGATGATGATGATGATGATGATGATGATGATGAT 64
Qy 1642 CAGCATGCTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1701
Db 65 CCGCATGGTGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 124
Qy 1702 AATTTAAACCTGAGAGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1761
Db 125 AGTCAAGGCGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 181
Qy 1762 CGGATTTGAGACTTCACCTTTGCGGTGCGGTGCTGATGATGATGATGATGATGATGATGAT 1821
Db 182 CGGATCTCGGCTCGCGCTTTGCGGTGCGGAGCGGA--GTGCCCGGCAAGTCATCG 240
Qy 1822 GTTTTACTACCTTATGTTTGAAGCGGATGATG 1856
Db 241 CGGTGCTACCGTGGATTTCTGGTGGCCACCTTG 275

RESULT 6
US-08-948-564-9
Sequence 9, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminezy, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESSES:
ADDRESSER: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512ch Carolina
COUNTRY: USA

```

/ ZIP: 27627
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/948,564
/ FILING DATE:
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Bennett, Virginia C.
/ REGISTRATION NUMBER: 37,092
/ REFERENCE/DOCKET NUMBER: 5051-409
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-854-1400
/ TELEFAX: 919-854-1401
/ INFORMATION FOR SEQ ID NO: 9:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1788 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 6..1601
/ US-08-948-564-9

Query Match 4.4%; Score 89; DB 3; Length 1788;
Best Local Similarity 48.2%; Pred. No. 8e-15;
Matches 285; Conservative 0; Mismatches 300; Indels 6; Gaps 1;

QY 1356 CAGGAGATGATATTCAAGGAAACACACAGTTGCGCTTAAATGAGTATCTGCT 1415
DB 972 CTGATCTGATTTTACGAGGAGGACACCCATGATTCATTAATGAGTATCTGCT 1031
QY 1416 AGGATGCTTCAATCCAGATATGCAATCAACGGTATCAAAAGAGTATCAAGATC 1475
DB 1032 CTGCTACTTAACATCAATGAACTAAAGATCAAGATGATGACATTTATTT 1091
QY 1476 GGGAAATCAAGAGCCCTAGATGAATCTGATGCTTCACTTCATATCTAACGCTGTG 1535
DB 1092 GGGAAAGACAGAGAGTGAAGATCTGACATTAACAAAGTGTGATCTCCAGCCATT 1151
QY 1536 GTGAAGAAGATGATGAGGCTTCATCCCGAGCCCACTTCTATCATGAGGCCGTTGGCC 1595
DB 1152 GTGAAGGAAACAAATGCGGCTATATCCACCAAGTCTTTATCACTTGTGAGCCATG 1211
QY 1596 ATTAACAGACGATGCTGATGATGCTGCTGCTGCGAGGAGACACAGATGTGAAC 1655
DB 1212 GAAGCTGACACCTTCTCAGGTGCTATCAATCTCTGCGACACGTTAAATGTGAAT 1271
QY 1656 ATGAGGCGCGATGCGATGATCAACAGTGTGAGTATCTTGAAGTTAAAGCTGAG 1715
DB 1272 GCTTGAAGATCCACCGGAGTGTGTGTTGAGATATCTCTATGATTTCAAGCTGGA 1331
QY 1716 AGTTCGTGCGAAAGAGGTGAGTGAAGTTTCGATCTTGGTGTGATTTGAAGACTT 1775
DB 1332 AGTTCGTGCGAAAGAGGTGAGTGAAGTTTCGATCTTGGTGTGATTTGAAGACTT 1385
QY 1776 GCACCTTCGCGGTGCGATGATTTGACCGGAGAAAGATCTGTTTATCAACGCTT 1835
DB 1386 GTTCCTTTGTTGTTGAGAGAGAGATGCTTGAAGCTTGGTGTGCTGTGCTGTG 1445
QY 1836 ATGTTTGAAGCGCGATGATTTACATGATGTTGAATGAGAGCCGTCGATGTAACGC 1895
DB 1446 CACTTGACCATGCGTGAATGATTTCAATTTCTTCAATGTCTTCTTCAATCAAGTT 1505
QY 1896 GTTGACTTATCTGAGAACTGAGGCTTCTTGCGAGATGCTTATCTCTT 1946
DB 1506 GTTGAATGACAGAGAGCATGACTCAAAATTTAAAGCAACCCGCTT 1556

```

```

RESULT 7
US-08-532-065B-1
/ Sequence 1, Application US/08532065B
/ Patent No. 5753507
/ GENERAL INFORMATION:
/ APPLICANT: Ohta, Daiisaku
/ APPLICANT: Mizutani, Masaharu
/ TITLE OF INVENTION: Plant Germplasm/Herol 10-Hydroxy]ase and
/ TITLE OF INVENTION: DNA Coding Therefor
/ NUMBER OF SEQUENCES: 6
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: No. 5753507artis Corporation
/ STREET: 59 Route 10
/ CITY: East Hanover
/ STATE: NJ
/ COUNTRY: USA
/ ZIP: 07936
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/532,065B
/ FILING DATE: 22-SEP-1995
/ CLASSIFICATION: 800
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Meigs, J. Timothy
/ REGISTRATION NUMBER: 38,241
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 919-541-8587
/ TELEFAX: 919-541-8689
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 1893 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ ORIGINAL SOURCE:
/ ORGANISM: Arabidopsis thaliana
/ IMMEDIATE SOURCE:
/ CLONE: P450-4
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 6..1490
/ US-08-532-065B-1

Query Match 4.4%; Score 88.8; DB 1; Length 1893;
Best Local Similarity 50.5%; Pred. No. 9.4e-15;
Matches 278; Conservative 0; Mismatches 257; Indels 15; Gaps 2;

QY 1340 TTTTGAATTTATTTACAGAGATGATATTCAAGGAAACACAGATGCGCTTAT 1399
DB 905 TATTGAACACTTCTCTGATATGTTTACAGAGACGATACAGATCTAGTACCT 964
QY 1400 CGAGTGAATCTCTGATGATGATGCTCTTCAATCAATATGCAATCAACGGTACAAAGCA 1459
DB 965 GGAATGCGCAATGACAGAGTATCTTAAACCTTAAAGATGCGCAAGCTCAGGCCA 1024
QY 1460 GCTGATCAAGTATGTCGGAATCAAGAGCCCTAGATGATGATGATGATGATGATGAT 1519
DB 1025 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1084
QY 1520 ATATCTAACGCTGTGTGTAAGAGATTTGAGCTTATCTCTCCAGGCCCACTTATC 1579
DB 1085 GTATTTCAGAGAGTGTGTAAGAGATTTGAGCTTATCTCTCCAGGCCCACTTATC 1144
QY 1580 ATGAGCCGCTTGGCCATTAACAGACGATGTTGATGATGATGATGATGATGATGATGAT 1639
DB 1145 CCGG---GAAAGCGGATCTGATGCGAGATTTCTGTTTCATGATGCTTAAAGATAC 1201

```

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rml

Page 6

QY 1640 CACAGCAATGTGAACATGCGGCGGTATGCGATGATCCACAGCGTGGTGTGATCCTTT 1699
Db 1202 TCAGGTTCTAGTGAAGCTCTGGGCCATAGAGAGACCCGAGCGTGTGGATATCCGTC 1261
QY 1700 GAGGTTTAACTGAGAGGTTCTGTGCAAAAAGAGTGAAGTGTGATTTGGTCTTGG 1759
Db 1262 CCGATTGAGCCGAGAGGTTTGGGAGAAATATGACGAGAGGTAGATATGAGA 1321
QY 1760 GTGCGATTGAGACTGACACTTCCGGTGGGTGCGTGGATTTGCCCCGGGAAATCT 1819
Db 1322 G-----CTTACACANTGGCGCGCGAGAGTGAATTTGCCCGGGAATGCCCTT 1369
QY 1820 TGGTTTACTACCGGTATGTTTGGACGGCGATGATGTTACATGATGTTGAATGGGACC 1879
Db 1370 GGGTATGAAAACAGTCTCTTATGCTTCTCTCTTATTCCTTTGACTGGAAGCT 1429
QY 1880 GTCCGATGCT 1889
Db 1430 TCCGAGGCT 1439

RESULT 8
US-08-606-505B-63
Sequence 63, Application US/08606505B
Patent No. 6114601

GENERAL INFORMATION:
APPLICANT: KIRUCHI, Yasuhiro
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: SHIMADA, Shigeto
APPLICANT: OHMAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasuhiro
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/V
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992

ATTORNEY/AGENT INFORMATION:
NAME: Peiray, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2100
INFORMATION FOR SEQ ID NO: 63 :
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
US-08-606-505B-63

Query Match 4.2% Score 84.6; DB 3; Length 2174;
Best Local Similarity 49.4%; Pred. No. 1,58-11;
Matches 306; Conservative 0; Mismatches 304; Indels 9; Gaps 3;

QY 1338 ATTTTGAATTTATTTACAGAGATGATTTACAGAGAACAGACAGATTTGGCTTTA 1397
Db 971 AATATCAAGGCTCTTTTATTAACATGTTTACCGCTGTAGATATCATCTTCAAGGCTC 1030
QY 1398 ATCGAGTATCTCGCTAGAGATGATGCTCTTATCATAGATATGCAATCAAGGCTA 1457
Db 1031 ATAGAGTGGCGCTGCGAGATGTTGTAAGATTCATATCTTAACAGAGCCAGAA 1090
QY 1458 GAGCTGATCAAGTATGTTGGGAAATCAAGAGCCCTAGATGATGATGATGCTTCACTT 1517
Db 1091 GAAATGAGCGGTGATGATGAGCGGAGACCGGCGGTTCTTAGAGGACATATCAAGTTG 1150
QY 1518 CCATATCTAACGCTGTGTGTAAGAAAGATTTAGGCTTATCCTCCAGGCCCACTTCTA 1577
Db 1151 CCATATCTCAAGCATCTGCAAGAGCTTTCAAGAGATCTCTTCAAGGCTTTA 1207
QY 1578 TCATGGGCGGCTTTGGCCATTAACAGACGATGTTGATGTCCTTCTTCCGACGGG 1637
Db 1208 AATCTCCACGAAATGCGCTCCAGCATGTAAGTAATGACATACATACCAAGAGGCT 1267
QY 1638 ACCAGCAATGTGAACATGAGGCGGTATGCGATGATCCACAGTGTGGTGTGATCT 1697
Db 1268 ACTAGGCTCAGGCTTAATATATGAGCTATTTGGAAGATATCATCTGTGTGGAAATCCA 1327
QY 1698 TTGAGATTAACTGAGAGGTTGCTGGCAAAAGAGTGAAGTGTGATGTTTGGCTTT 1757
Db 1328 AATGATTTTAACTCGATAGTGTGTTTGGAAAGAAATGCA---AGATCGATCCACGA 1384
QY 1758 GGGTGGATTTGAGACTTGACACTTCCGGTGGGTGCGATGATTTCCCGGAGAAAT 1817
Db 1385 GGAATGATTTGAGCTATATCCATTTGAGACTGAGAGAAATTTGGCTGGAAACAGA 1444
QY 1818 CTGGTTTACTACCGGTATGTTTGGACGGCGATGATGTTACATGATGTTGAATGGGA 1877
Db 1445 TTGGAAATCTTCTAGTGAAGATATTTTGGAACTTTGGTGAATCTTTTGTGGAA 1504
QY 1878 ---CCGTCCGATGTAAAGGCGGTGACTTATCTGAGAACTGAGGCTTTCTTGGAGAT 1934
Db 1505 TTGCATCTCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1564
QY 1935 GCTAATCTCTTCTCTGCTTA 1953
Db 1565 GCGAGGCTCTTCTGCTGCTTA 1583

RESULT 9
US-09-616-990-63
Sequence 63, Application US/09616990
Patent No. 6232109

GENERAL INFORMATION:
APPLICANT: KIRUCHI, Yasuhiro
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: SHIMADA, Shigeto
APPLICANT: OHMAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasuhiro
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/V
OPERATING SYSTEM: MS-DOS Ver3.30

SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-JUL-2000
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perity, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 63 :
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 63
US-09-616-990-63

Query Match 4.2% Score 84.6; DB 3; Length 2174;
Best Local Similarity 49.4%; Pred. No. 1.5e-13;
Matches 306; Conservative 0; Mismatches 304; Indels 9; Gaps 3;
QY 1338 ATTTTGAATTATTATTACAGAGATGATATTCAGAGAAACAGACACAGTGGCGTCTTA 1397
DB 971 AATATCAAGGCTCTTTATTGAACATGTTTACCGGTACGATCATCTTCAACCGTC 1030
QY 1398 ATCGAGTGGATCCCTCGTAGAGTGCCTTCATCCAGATGATGCAATCAACGGTCAAAAAC 1457
DB 1031 ATAGAGTGGCGCTGCGCGAGGTGCTAAAGATTCATCATCTTAAGACAGGCCCAAGAA 1090
QY 1458 GAGCTGATCAAGTAGTCCGGGAAATCAAGAGCCCTAGATGATCTGACTTGCGTCACTT 1517
DB 1091 GAATGACGGGTGATCGGCGAGACCGCGGTTCTTGAAGCAGCATATCAAAAGTTG 1150
QY 1518 CCATATCTAACCGCTGTGTGGAAGAGATTTAGGCTTATCTCCAGGCCCACTTTCA 1577
DB 1151 CCAATCTCCAAAGCATCTGCAGAAAGAGCTTTGAGAAAGATCTCTT--CAGCGCTTTA 1207
QY 1578 TCATGGGCGCGTTGGCCATTAACAGACACAGATCGTTGATGTCGTTGTTCCGGAGGG 1637
DB 1208 AATCTCCAGGAATCGGTGCGCAAGCATGGAAGTAATGGAACATCAATCAACAAAGGGC 1267
QY 1638 ACCACAGCATGTGAACATGTGGCCGCTATCGATATCAACGCTGTGGGTGATCTT 1697
DB 1268 ACTGAGCTCAGCGCTTAACATATGCGCTATGGAAGAGATCCATCTGTGTGGGAAATCCA 1327
QY 1698 TTGAGTTTAAACCTGAGAGGTTGCTGCGCAAAAGAGGTGAGGAGTTTGGTCTT 1757
DB 1328 AATGAATTTAACCTGTATAGTCTTTTGAAGAGAAAGATGCA---AGATGATTCACGA 1384
QY 1758 GGGTCGATTTGAACATCTTGACCTTTGGGCTCGGCTGTGGATTTGCCCGGAGAAAT 1817
DB 1385 GGAATGATTTTGAAGTATCGCATTTGAGCTGGAAGAGAAATTTGCGGTGAACAAAG 1444
QY 1818 CTGTGTTTATACCGTATGTTTGAACGCGGATGATGATCACTGATGTTGATGGGA 1877
DB 1445 TTGGGAATCTTCTAGTGAATATATTTTGGAACTTTGTGCACTCTTTTGTGGGA 1504
QY 1878 ---CCGTCGATGTGAACGCGGTGACTTATCTGAAGAACTGAGGCTTTCTTGGAGATG 1934
DB 1505 TTGCATCTCTCTGATGATGAACATTAACATGATGATGATCTTTGGGCTTGTCTGCAAGAG 1564

QY 1935 GCTATCTCTCTGCTA 1953
DB 1565 GCAGTGCCTCTGTGCTA 1583

RESULT 10
US-09-033-055A-3
Sequence 3, Application US/09033055A
Patent No. 6069241
GENERAL INFORMATION:
APPLICANT: OKAWA, HIDEO
APPLICANT: IMASHI, HIROMASA
TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,055A
FILING DATE: MARCH 2, 1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PAUL E. WHITE, JR.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 9437/251563
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3651
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1761
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-033-055A-3

Query Match 4.1% Score 83.2; DB 3; Length 1761;
Best Local Similarity 49.0%; Pred. No. 3.3e-13;
Matches 282; Conservative 0; Mismatches 280; Indels 6; Gaps 2;

QY 1351 TATTACAGAGATGATATTCAGAGAACAGACACAGTGGCGTCTTAATGAGTGAATCC 1410
DB 962 TAATCAGATATGATATCTGACAGCACAGATACATCTGCTGTCACCAACGAATGGCCAA 1021
QY 1411 TGGCTAGATGCTCTTATCCAGATATGACATCAACGCTACCAAGAGATCTGATATG 1470
DB 1022 TGGCTAGATGATTAAGATCCAAATGCTCTCAAGAGATCCAAAGAGATCTGATATG 1081
QY 1471 TAGTGGGAAATCAAGAGCCCTAGATGATCTGACTTGGCTCACTTCAATATCTAACG 1530
DB 1082 TTGTTGATTCGACCGGTATGGAACGATCCGATCGATCTGTTCAATCAAGTACCTTGGT 1141
QY 1531 CTGTGTGAAGAGATATTTAGAGCTTCACTCCAGAGCCCACTTATCATGGGCGGT 1590
DB 1142 GTGTAGTGAATCAACATTTGAAATGACACCTGCTGCTCACTTTCT--AATTCACATG 1198
QY 1591 TGGCATTAACAGACAGATCGTTGATGCTGCTTGTTCGGCAGAGGACCAAGCAATGG 1650
DB 1199 AATCAATTCAGATATCAAGATCAACGGGTATATATCCAGCAAGACAGCTGTCTTCA 1258
QY 1651 TGAACATGTGGCCGCTATGCAATGATCCACAGCTGTGGTATCTTGAATTTAAAC 1710
DB 1259 TCAACACATGCTTGTGTGAAGAAACAAAGATATGGAATATGATGATGATTTAGGC 1318

QY 1711 CTGAGAG--GTCGTGGCAAAAGAGGTGAGTGGAGTTTGGGTTCTTGGGCGGATT 1767
DB 1319 CAGAGAGACATTTGGCTTGCATGATGAAATTAAGTGAATGATGATGAGGCGGATT 1378
QY 1768 TGAAGCTTGACCTTTCGGGTCGGGTCGTGCGATTGCCCCGGAGAAATCTTGGTTTA 1827
DB 1379 TCAAGATTTTACCAATTTAGTGTCTGGAAGAGTGTCTGCGACCAATTAGAGTGA 1438
QY 1828 CTACCGTTATGTTTGGACGGCGAGTGAATGTTACATGATTTGATGGGACCGTCGATG 1887
DB 1439 AATTGGTACTTATGAGCTTTGGCTGATGTTGCTTGCCTTGTGAGTGGAGCCACAGATG 1498
QY 1888 GTAACGGCGCTTGAATCTGAGAACTGAGGCTTT 1923
DB 1499 GATTACGGCTGAGATATGACACAAATTGAGGTTT 1534

RESULT 11

US-08-606-505B-64
Sequence 64, Application US/08606505B
Patent No. 6114601

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIYOKAWA, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:

ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/IV
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606.505B
FILING DATE: 23-MAR-1996

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear

MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Campanula medium
FEATURE:

NAME/KEY: CDS
LOCATION: 180 to 1748
IDENTIFICATION METHOD: by experiment
US-08-606-505B-64

Query Match 4.1%; Score 83.2; DB 3; Length 1927;
Best Local Similarity 49.5%; Pred. No. 3.4e-13;
Matches 273; Conservative 0; Mismatches 273; Indels 6; Gaps 2;

QY 1337 AATTTTGGATTAATTATTAACAGAGATGATATTCAGAGAAACAGACAGTTGGGCTT 1396
DB 1097 AATGTTAAGGACATCTCTTTTGGATTTATTCACGGCGGGGCAOGGATCATCATGAAGT 1156
QY 1397 AATCGATGAGATCTTCCTGAGGATGCTCTTCATCATGATATTCAGATCAAGATACAAA 1456
DB 1157 GATCGAATGGGCACTGACCGAATGTTAAGCATGACAGATCTTAAACGGGGCCACGA 1216
QY 1457 CGAGCTGATCAAGATGTCGGGAATCAAGACCCCTAGATGATCTGACTTGGCTTCACT 1516
DB 1217 AAAAAAGACAGATGATTTGGGAGAAAGAAAGAACTAAGATCTGACATACAAAATT 1276
QY 1517 TCCATATCTAACGGCTGTGTGGAAGAAAGATTTGAGGCTTCATCTCCAGGCCACTTCT 1576
DB 1277 GCCATATTTCCAGGCCATATGCAAGAAACATTCGAAACACCCCTTCC--CAGGCCCTT 1333
QY 1577 ATCATGGGCGCGTTTGGCCATTAACAGACAGATCGTTGATGTCGCTTGTTCGGCAGG 1636
DB 1334 AATCTCTCAAGAAATCTCAACAGAGCATGTGAATGACGATTTCAATACAAAATA 1393
QY 1637 GACCAAGCAATGTGAACATGTGGCCGTATCGCATATCCACGTTGGTGTATCC 1696
DB 1394 CACTAGACTAATAGAACATATGGGCAATAGGGAGGAGACCTTAAGTGTGGAAATCC 1453
QY 1697 TTTGAGTTTAACTGAGAGGTTGTGGCAAAAGAGTGAAGTGGAGTTTGGTTCT 1756
DB 1454 ATTAGTTTACTCGGACGTTTCTTGAGTGAAGAAC--ACGGAAATATGATCCGCG 1510
QY 1757 TGGGTGAGTTTGAAGCTTGACCTTTCGGGTGGGTGTCGGAATTTGCCCGGAGAA 1816
DB 1511 AGGTATATCAATTTGATTAATCCATTGGGGCTGACGAAAGATATGACAGGGGCTAG 1570
QY 1817 TCTGTTTACTACCGTTATGTTTGGACGCGCATGATTTACATGATTTGAATGGGG 1876
DB 1571 AATGGAGGGCGCTCGGTGAGTACATTTAGTACATTTGGTGCCTATTTGATTGAA 1630
QY 1877 ACCGTCGATGG 1888
DB 1631 ATTGCTGATGG 1642

RESULT 12.

US-09-616-990-64
Sequence 64, Application US/09616990
Patent No. 6232109

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIYOKAWA, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi

TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/IV
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616.990
FILING DATE: 14-Jul-2000

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992

Page 9

```

; TITLE OF INVENTION: A Method For Regulation Of Plant Nighn Composition
; FILE REFERENCE: 7024-325
; CURRENT APPLICATION NUMBER: US/09/091.432
; CURRENT FILING DATE: 1998-06-18
; EARLIER APPLICATION NUMBER: PCT/US96/20094
; EARLIER FILING DATE: 1996-12-19
; EARLIER APPLICATION NUMBER: US 60/009,119
; EARLIER FILING DATE: 1995-12-22
; EARLIER APPLICATION NUMBER: US 60/013,388
; EARLIER FILING DATE: 1996-03-14
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: Microsoft Word 2.0C
; SEQ ID NO 3
; LENGTH: 5156
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-091-432-3

      Overy Match          4.1%; Score 81.4; DB 2; Length 5156;
      Best Local Similarity 49.6%; Pred. No. 1,7e-12;
      Matches 276; Conservative 0; Mismatches 266; Indels 15; Gaps 2;

```

QY	1334	TAATAATTTTGAATTATTAATTAACAGAGAGATATTACAGAGAAACACACACAGTTGGGT	1353
Db	3877	TGATTTTTTATTTGATTTATTAGACGTATCTTTGGAGGAAACGGAACGGTAGCGTC	3936
QY	1394	CTTATATGAGATGATCTCTCCCTAGAGACGCTTCATCCAGATATGATATCAACGGGTACA	1453
Db	3937	GGCGATAGAGGGGCCCTTAACGAGATTATTAACGAGCCCCGGAGATCTTAACGGGTCCA	3996
QY	1454	AAACGAGCTGATCAAGTACAGTACGGGAAATCAAGACCCCTAGATGATCTGACTTGGCTTC	1513
Db	3997	ACAAGAACCTGCCAGATCGTTGGACTTGCACAGACGAGTTGAGAAATCCGACATCGAGAA	4056
QY	1514	ACTTCATATCTAACGGCTGTGGTGAAGAAGTATTAGGCTTCATCCTCCAGGCCACT	1573
Db	4057	GTTTACTTATCTCAATATGCACATCAAGAAACCCCTAAGATGACCCACCGATCCCTCT	4116
QY	1574	TCATATCATTGGCCCCGTTTGGCCATACAGACACGATGTTGATGTCGTCTGTTCCGGC	1633
Db	4117	CCT-----CCTCCACGAAACCCGCGAGGACACTAGTATGACGCGTTTCTCATTTCCCA	4170
QY	1634	AGGACACACAGCATATGTTGAACATGTGGGCCGTATCCATGATCCACACGTTGGGTGTA	1693
Db	4171	GAATCTCTGTATGATGATCAACGGCTTTGCCATATAGACGGGACCCACCTCTTGGACATGA	4230
QY	1694	TCCTTTGAAGTTTAAACCTGAGAGGTTCTGTGGCAAAAGAAAGTGAAGTTTGGGT	1753
Db	4231	CCCGACACGTTTAAAGCATCGAGGTTTTTGGAAACCGGCGTACCAG-----ATT	4281
QY	1754	TCTTGGCTCGGATTGAGACTTGACACCTTTCGGGTCGGGTCTGTGCAATTGGCCCCGGGAA	1813
Db	4282	CAAGGGAGCATTTTCAGATTATACCGTTTCGGGTCCGGTGTGATCTGTCCCGGGAT	4341
QY	1814	GAATCTTGTTTACTACCGTTATGTTTGGACGGGCATATGTTAATAGATTTGATG	1873
Db	4342	GCAACTGAGGGTTATACCGGCTGACTTAAGCCGTGGCTCATATTAATTAATTCCTTCACGTG	4401
QY	1874	GGGACCGTCCGATGTA	1890
Db	4402	GAATTAACCTGATGGA	4418

```

RESULT 14
US-09-387-663-3
; Sequence 3, Application US/09387663
; Patent No. 6489538
; GENERAL INFORMATION:
; APPLICANT: Chapple, Clint
; TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
; FILE REFERENCE: 7024-325
; CURRENT APPLICATION NUMBER: US/09/387,663
; CURRENT FILING DATE: 1999-08-31

```

PRIOR APPLICATION NUMBER: US 09/091,432
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: PCT/US96/20094
PRIOR FILING DATE: 1996-12-19
PRIOR APPLICATION NUMBER: US 60/009,119
PRIOR FILING DATE: 1995-12-22
PRIOR APPLICATION NUMBER: US 60/013,388
PRIOR FILING DATE: 1996-03-14
NUMBER OF SEQ ID NOS: 3
SOFTWARE: Microsoft Word 2.0C
SEQ ID NO 3
LENGTH: 5156
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-387-663-3

Query Match 4.1%; Score 81.4; DB 4; Length 5156;
Best Local Similarity 49.6%; Pred. No. 1.7e-12;
Matches 276; Conservative 0; Mismatches 266; Indels 15; Gaps 2;

QY 1334 TAAATTTTGAAGTATTATTAACGAGATGATATTCAGAGAACAGACAGTTGGGT 1393
DB 3877 TGATTTTATTTGATTTTATTTAGACGTTATGTTGAGAGAACGAAACGTAAGCTC 3936
QY 1394 CTTATCGAGTGAATCTCGTAGAGATGCTCTTCATCAGATATGCAATCAACGGTACA 1453
DB 3937 GGGCATAGAGTGGGCTTAACGAGATATTAACGAGCCCGAGAGATCTAAACGGGTCCA 3996
QY 1454 AACAGAGTGAATCAAGTAGTGGGAAATCAAGAGCCCTAAGTAAATGACTGGCTTC 1513
DB 3997 ACMAAGACTCGCCGAGAGTGTGAGCTGACAGAGAGTTGAAGATCGACATGAGAA 4056
QY 1514 ACTTCATATCTAAGCGGTGTGTAAGAAAGATATGAGGCTTCATCTCCAGGCCACT 1573
DB 4057 GTTGAATATCTCAATGACACTCAAGAAACCTTAAGATGACCCAGATCCCTCT 4116
QY 1574 TCTATCAGTGGCCCGTTGGCCATTAACAGACAGATGTTGATGCTCTGTTCCGGC 1633
DB 4117 CTT-----CTTCACAGAAACCGGAGAGACATGATACAGGTTTCTTCATTTCCAA 4170
QY 1634 AGGAGACACAGAGATGTAACATGAGGCGCTATGCAATGATCCACAGTGGGTGA 1693
DB 4171 GAAATTCGTGTGATATCAACGCGTTTCCATAGAGCCGACCAACCTCTTGGACTGA 4230
QY 1694 TCTTTGAGATTAAACCTGAGAGTTCTGTGCAAAAGAGGTGAGTGGATTTTGGCT 1733
DB 4231 CCGGACAGTTTAAACATCAAGGTTTGGAAACCGGCGCTAACCGG-----ATTT 4281
QY 1754 TCTTGGGTGGATTGAGACTGACCTTGGGGTGGGGTGGTGGGATTTGGCCGGGAA 1813
DB 4282 CAAGGAGACATTTGAGATTATACCGTTCCGGTGGGTCTTAATGCTGCGCGGAT 4341
QY 1814 GAATCTGTTTACTACCGTTATGTTTGGACGGGATGATGTTAAGTTGAATG 1873
DB 4342 GCACTAGGTTATACCGCTTGAATGACCGGCTCATATATTAACATGCTTACGCTG 4401
QY 1874 GGAACCGTCCGATGTA 1890
DB 4402 GAAATTAACGATGGA 4418

RESULT 15

US-09-033-055a-1
Sequence 1, Application US/09033055A

Patent No. 6068241
GENERAL INFORMATION:
APPLICANT: OHKAWA, HIDEO
APPLICANT: IMAISHI, HIROMASA
TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.

CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,055A
FILING DATE: MARCH 2, 1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PAUL E. WHITE, JR.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 9437/251563
TELEPHONE: (202) 861-3651
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1722
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-033-055a-1

Query Match 3.9%; Score 79.2; DB 3; Length 1722;
Best Local Similarity 49.5%; Pred. No. 4.2e-12;
Matches 275; Conservative 0; Mismatches 263; Indels 18; Gaps 2;

QY 1319 AAAAGTAAAGTACTTAAATTTTGAATTTATTAACAGAGATGATTTCAAGAGAAC 1378
DB 849 ACATGCAATTTAGTAAATCAATATAGAGGCTTGTGATGACCTTTTATGCGGGAC 908
QY 1379 AGACAGATGGGGCTTATATGAGTGAATCTCGCTAGAGATGCTTCATCAGATAT 1438
DB 909 TGATTAAGTTCAAAATACATTAAGATGGGCAATGTTAGCAATAGAGAAACATACAT 968
QY 1439 GCAATCAACGGTACAAACGAGCTGATCAAGTAGTGGGAAATCAAGAGCCCTAGATGA 1498
DB 969 AATGAAAAAGCTAAACGAGCTACAGAAATTTAGCAAGGAGGCTAATATAGAA 1028
QY 1499 ATCTGACTTGGCTTCACTTCCATATCTAACGGGTGTGTAAGAGATGAGGCTTCA 1558
DB 1029 AGCTGATTTGACGCTCTCCATATTTGCAATGATTTTAAAGAAACCTTAAGATGCA 1088
QY 1559 TCTTCAGGCCCACTTATCATGAGGCGGTTGGCCATTAACAGACAGATGTTGATGG 1618
DB 1089 CCGACAGGCTCCCTTTT-----ATTGCAAGTTGATCAAGAGCTTGAGGCGTGTGG 1142
QY 1619 TCGTCTTGTTCGGCAGAGACACAGCATGTTGAACATGTTGGCCGTATCGCATGATCC 1678
DB 1143 CTATTTTGTCCAAAACACTCAAGTGTGTGCAATATGTCATATAGCCGCAACC 1202
QY 1679 ACAGTGTGGATTGATCTTTGAGATTAAACCTGAAGAGTTGCTGGCAAAAGAGTGA 1738
DB 1203 AACTATATGGAGATCTTTTGGATTTAAGCCGAAAGATTTGGGGGTCAAAAAATGGA 1262
QY 1739 GGTGGAATTTTGGTTCTTTGGGTCCGATTTGAAGTATGCACTTTGGGGTCCGAGTGG 1798
DB 1263 C-----GTTGTGACAAAGATTTGAATCATTCATTTGCTGTCGAGCAAG 1310
QY 1799 GATTTGCCCGGAGAGATCTTTGTTTACTACCGTTATGTTTGAACGGCGATGATGTT 1858
DB 1311 AATTTGCCCGGATTTACTTTGGCAACAGACATTAACGAATGTTGGTTTCAATGTT 1370
QY 1859 ACATGAGTTGAATGG 1874
DB 1371 GAATTCATTTGATTTGG 1386

Tue Jan 20 17:43:17 2004

us-10-022-025a-1.rnt

Page 11

Search completed: January 16, 2004, 13:43:40
Job time : 129 secs

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rmpb

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 03:16:41 ; Search time 538 Seconds
(without alignments)
10436.706 Million cell updates/sec

Title: US-10-022-025a-2
Perfect score: 1593
Sequence: 1 atggctacgaactcgaag.....tgcgcgctagcgagcttaa 1593

Scoring table:
Gapop 10.0, Gapext 1.0

Searched: 2324096 seqs, 1762381658 residues

Total number of hits satisfying chosen parameters: 4648192

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:*

1:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
2:	/cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
3:	/cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
4:	/cgn2_6/prodata/1/pubpna/US06_PUBCOMB.seq:*
5:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
6:	/cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
7:	/cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
8:	/cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
9:	/cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
10:	/cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
11:	/cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
12:	/cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
13:	/cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
14:	/cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
15:	/cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
16:	/cgn2_6/prodata/1/pubpna/US10_PUBCOMB.seq:*
17:	/cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*
18:	/cgn2_6/prodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1593	100.0	1593	US-09-349-385-5	Sequence 5, Appl1
2	1593	100.0	1593	US-10-022-025a-2	Sequence 2, Appl1
3	1360	85.4	2009	US-10-022-025a-1	Sequence 1, Appl1
4	976	61.3	1902	US-09-349-385-2	Sequence 2, Appl1
5	794	49.8	7869	US-09-349-385-3	Sequence 3, Appl1
6	651.6	40.9	1611	US-09-349-385-7	Sequence 7, Appl1
7	599.6	37.6	1764	US-10-163-198-20	Sequence 20, Appl1
8	585.8	36.8	1905	US-10-163-198-24	Sequence 24, Appl1
9	355	22.3	1799	US-09-349-385-13	Sequence 13, Appl1
10	352	22.1	1930	US-10-163-198-34	Sequence 34, Appl1
11	348.2	21.9	1990	US-09-349-385-9	Sequence 9, Appl1
12	316.6	19.9	1934	US-10-163-198-22	Sequence 22, Appl1
13	290.6	18.2	1687	US-10-163-198-40	Sequence 40, Appl1
14	286	18.0	1545	US-10-163-198-96	Sequence 96, Appl1
15	283	17.8	1758	US-10-163-198-94	Sequence 94, Appl1

16	278.2	17.5	1597	15	US-10-163-198-16	Sequence 16, Appl1
17	275.2	17.3	1585	15	US-10-163-198-92	Sequence 92, Appl1
18	274.4	17.2	1539	15	US-10-163-198-18	Sequence 18, Appl1
19	265.8	16.7	2924	15	US-10-163-198-26	Sequence 26, Appl1
20	237.4	14.9	2087	10	US-09-349-385-11	Sequence 11, Appl1
21	234	14.7	1146	15	US-10-163-198-12	Sequence 12, Appl1
22	233.4	14.7	1578	15	US-10-163-198-12	Sequence 1, Appl1
23	213.4	13.4	1121	15	US-10-163-198-8	Sequence 8, Appl1
24	194.8	12.2	1357	15	US-10-163-198-36	Sequence 98, Appl1
25	189.2	11.9	884	15	US-10-163-198-38	Sequence 38, Appl1
26	156.4	9.8	442	15	US-10-163-198-38	Sequence 3, Appl1
27	155.6	9.8	1720	15	US-10-163-198-3	Sequence 1, Appl1
28	129.6	8.1	1708	9	US-09-796-256A-1	Sequence 19, Appl1
29	129.2	8.1	1721	12	US-10-429-949-19	Sequence 14, Appl1
30	125	7.9	178	15	US-10-163-198-14	Sequence 1, Appl1
31	119.4	7.5	1527	9	US-09-931-267-1	Sequence 1, Appl1
32	119.4	7.5	1549	9	US-09-931-267-2	Sequence 2, Appl1
33	117.8	7.4	1527	9	US-09-931-267-2	Sequence 42, Appl1
34	111.2	7.0	1673	12	US-10-431-273-42	Sequence 7, Appl1
35	111.2	7.0	1714	12	US-10-431-273-43	Sequence 43, Appl1
36	106.8	6.7	1770	12	US-10-411-115-7	Sequence 1, Appl1
37	106.8	6.7	1933	14	US-10-067-534-1	Sequence 30, Appl1
38	105.6	6.6	1457	15	US-10-163-198-30	Sequence 40, Appl1
39	104.8	6.6	1747	12	US-10-431-273-40	Sequence 2669, Appl1
40	104.8	6.6	1763	12	US-10-431-273-41	Sequence 41, Appl1
41	101.6	6.4	1557	10	US-09-938-842A-2669	Sequence 55, Appl1
42	99.6	6.3	1545	13	US-10-356-151-55	Sequence 28, Appl1
43	99.6	6.3	1545	15	US-10-142-231-55	Sequence 3, Appl1
44	98.8	6.2	528	15	US-10-163-198-28	
45	97.2	6.1	1764	10	US-09-947-027-3	

ALIGNMENTS

RESULT 1
US-09-349-385-5
Sequence 5, Application US/09349385
Patent No. US20020152495A1
GENERAL INFORMATION:
APPLICANT: Ito, Toshio
APPLICANT: Meyersowitz, Elliot
TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
FILE REFERENCE: NBI-0002
CURRENT APPLICATION NUMBER: US/09/349,385
CURRENT FILING DATE: 1999-07-09
EARLIER APPLICATION NUMBER: 60/115,967
EARLIER FILING DATE: 1999-01-15
NUMBER OF SEQ ID NOS: 13
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 1593
TYPE: DNA
ORGANISM: Arabidopsis thaliana
FEATURE:
OTHER INFORMATION: CDNA
US-09-349-385-5

Query Match 100.0%; Score 1593; DB 10; Length 1593;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	ATGGCTACGAAGCTTCAATCTTGGCTTTCCTCCAAATGAGCGCTTCA	60
DB	1	ATGGCTACGAAGCTTCAATCTTGGCTTTCCTCCAAATGAGCGCTTCA	60
QY	61	AGCCAAACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	120
DB	61	AGCCAAACCACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	120
QY	121	CTCTCTTATGACCTATCCCGGTGACCTGCTGGGGGAATTAACCTTGGCGGGTAA	180

Db 121 CTCTTCTATGAGACCTATCCCGGTGAGACTGCTTG366GAAATACCTCTTCCGCGGTGA 180
 QY 181 ATATCCGGTTCATCAAAACCGGAAAGTTATTCGGGTCCAAAGGCTCCCTTGGTT 240
 Db 181 ATATCCGGTTCATCAAAACCGGAAAGTTATTCGGGTCCAAAGGCTCCCTTGGTT 240
 QY 241 GGAAGCATGTCTCATGTCAAGCACTCTAGCTCAGCAGAAATGCTGATGCACTGAG 300
 Db 241 GGAAGCATGTCTCATGTCAAGCACTCTAGCTCAGCAGAAATGCTGATGCACTGAG 300
 QY 301 AAATTCGAGCGAAGAGGCTCATGCTTTCAGCTTAAAGAGAGACTCCGCTGATGCTACG 360
 Db 301 AAATTCGAGCGAAGAGGCTCATGCTTTCAGCTTAAAGAGAGACTCCGCTGATGCTACG 360
 QY 361 TGCAATCCCGACGTAGCGAAGAGATTCGATATGAGCCCGGTTTTCGCTGATGCACTGAG 420
 Db 361 TGCAATCCCGACGTAGCGAAGAGATTCGATATGAGCCCGGTTTTCGCTGATGCACTGAG 420
 QY 421 AAAGATGCGCTTACTCTCATGTATGTTTAAACAGAGCAATGGTTTTCACACACAGGCTGT 480
 Db 421 AAAGATGCGCTTACTCTCATGTATGTTTAAACAGAGCAATGGTTTTCACACACAGGCTGT 480
 QY 481 TACTGCGAAGCGCTTCCGCTATGCTGATGCTGCAACCATCTCTTATGATCAAAACCAATGAG 540
 Db 481 TACTGCGAAGCGCTTCCGCTATGCTGATGCTGCAACCATCTCTTATGATCAAAACCAATGAG 540
 QY 541 AGAGCCGAGACGCAACGAGAGTATCAACCAAGATGTTGAGTTTCTTGAAGAACAG 600
 Db 541 AGAGCCGAGACGCAACGAGAGTATCAACCAAGATGTTGAGTTTCTTGAAGAACAG 600
 QY 601 AGTGTATCGAAGCCCTGTTTGTGCGAGAGTTGCTTAAACGCGGTGCTTAAACAATG 660
 Db 601 AGTGTATCGAAGCCCTGTTTGTGCGAGAGTTGCTTAAACGCGGTGCTTAAACAATG 660
 QY 661 ATGTGCTCTGTATTCGACCAAGATAGAGCTTGAAGAAACATGTTGAGTTACGTGAA 720
 Db 661 ATGTGCTCTGTATTCGACCAAGATAGAGCTTGAAGAAACATGTTGAGTTACGTGAA 720
 QY 721 ATGTGCTCTGTATTCGACCAAGATAGAGCTTGAAGAAACATGTTGAGTTACGTGAA 720
 Db 721 ATGTGCTCTGTATTCGACCAAGATAGAGCTTGAAGAAACATGTTGAGTTACGTGAA 720
 QY 781 CTATCGAGATTGATCTCAAGACTCCGCTCTAGATGTTTCAACCTGATCAACCAAGATA 840
 Db 781 CTATCGAGATTGATCTCAAGACTCCGCTCTAGATGTTTCAACCTGATCAACCAAGATA 840
 QY 841 AACCGGTTTGTATCCCGGATATATCCGCAACACCGTATCAACCGGATGTTTCCCTGCT 900
 Db 841 AACCGGTTTGTATCCCGGATATATCCGCAACACCGTATCAACCGGATGTTTCCCTGCT 900
 QY 901 GATTTGCTGAGAGTTTGTCTCTCCCTCATGAGTTTCAATTAATTAATCCGACCCGAGATA 960
 Db 901 GATTTGCTGAGAGTTTGTCTCTCCCTCATGAGTTTCAATTAATTAATCCGACCCGAGATA 960
 QY 961 ATGCGCGTTCTTTGGAGATGATATTCAGAGAAACAGACAGATTGCGGCTTATATGAG 1020
 Db 961 ATGCGCGTTCTTTGGAGATGATATTCAGAGAAACAGACAGATTGCGGCTTATATGAG 1020
 QY 1021 TGGATCTGCTAGAGATGCTTCTCATCCATGATATGATCAACGATCAACGATCAACGAGCTG 1080
 Db 1021 TGGATCTGCTAGAGATGCTTCTCATCCATGATATGATCAACGATCAACGATCAACGAGCTG 1080
 QY 1081 GATCAAGTATGTCGGAATCAAGAGCCCTAGATGATGATCTGATGCTTCACTTCCATAT 1140
 Db 1081 GATCAAGTATGTCGGAATCAAGAGCCCTAGATGATGATCTGATGCTTCACTTCCATAT 1140
 QY 1141 CTAAAGGCTGCTGTAAGAGATGATGAGGCTTCACTCCGACGCGCACTTCTATCATG 1200
 Db 1141 CTAAAGGCTGCTGTAAGAGATGATGAGGCTTCACTCCGACGCGCACTTCTATCATG 1200
 QY 1201 GCGCGTTTGGCATTAACAGACAGATGTTGATGCTGCTTGTTCGCGACAGGACACA 1260
 Db 1201 GCGCGTTTGGCATTAACAGACAGATGTTGATGCTGCTTGTTCGCGACAGGACACA 1260

QY 1261 GCAATGTGAAACATGTGAGGCGGTATGCGATGATCAACGATGAGTTGATCTTGGAG 1320
 Db 1261 GCAATGTGAAACATGTGAGGCGGTATGCGATGATCAACGATGAGTTGATCTTGGAG 1320
 QY 1321 TTTAACTGAGAGGTTTGTGCAAAAGAGTGAAGTGAAGTTTGTGAGTTTGGGTCG 1380
 Db 1321 TTTAACTGAGAGGTTTGTGCAAAAGAGTGAAGTGAAGTTTGTGAGTTTGGGTCG 1380
 QY 1381 GATTGAGACTTGCACCTTTCGGGTGCGGTGCTGAGATTTTCCCGGAAAGATCTTGGT 1440
 Db 1381 GATTGAGACTTGCACCTTTCGGGTGCGGTGCTGAGATTTTCCCGGAAAGATCTTGGT 1440
 QY 1441 TTTACTACCGTTATGTTTTCGACGCGGATGATGTTTCACTGAGTTTGAATGGGACCTTC 1500
 Db 1441 TTTACTACCGTTATGTTTTCGACGCGGATGATGTTTCACTGAGTTTGAATGGGACCTTC 1500
 QY 1501 GATGTAACGCGGCTTGAATTTCTGAGAACTGAGGCTTTCCTGCGAGATGCTATCT 1560
 Db 1501 GATGTAACGCGGCTTGAATTTCTGAGAACTGAGGCTTTCCTGCGAGATGCTATCT 1560
 QY 1561 CTTCCTGCTAAATTTGCGCGGTAGGCGGAGTTAA 1593
 Db 1561 CTTCCTGCTAAATTTGCGCGGTAGGCGGAGTTAA 1593
 RESULT 2
 US-10-022-025a-2
 ; Sequence 2, Application US/10022025A
 ; Publication No. US20030092014A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coleman, John R.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Polypeptides for Catabolism of Absc.
 ; FILE REFERENCE: 3310 0003
 ; CURRENT APPLICATION NUMBER: US/10/022,025A
 ; PRIOR FILING DATE: 2003-01-09
 ; PRIOR APPLICATION NUMBER: US 60/254,819
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 2
 ; LENGTH: 1593
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (1)..(1593)
 US-10-022-025a-2
 Query Match 100.0%; Score 1593; DB 15; Length 1593;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1593; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ATGGCTAACGAACTCGAAGCTCTTAATCTTGTGCTTGTGCAAAATGAGCGTTCTA 60
 Db 1 ATGGCTAACGAACTCGAAGCTCTTAATCTTGTGCTTGTGCAAAATGAGCGTTCTA 60
 QY 61 AGCCAAACCAACCTTCTCTCCCTCTGCGCGTCAACATCATCTGCTCCCATATCT 120
 Db 61 AGCCAAACCAACCTTCTCTCCCTCTGCGCGTCAACATCATCTGCTCCCATATCT 120
 QY 121 CTCTTCTATGAGACTATCCCGGTGAGACTGCTGGGGAATPAACTCTTGGCGGTTA 180
 Db 121 CTCTTCTATGAGACTATCCCGGTGAGACTGCTGGGGAATPAACTCTTGGCGGTTA 180
 QY 181 ATATCCGTTTCAATCAAAACCGGAAAGCTTATTCGCGTCAAAAGGCTTCCCTTGGT 240
 Db 181 ATATCCGTTTCAATCAAAACCGGAAAGCTTATTCGCGTCAAAAGGCTTCCCTTGGT 240
 QY 241 GGAAGCATGTCTCATGTGCAAGCACTCTAGCTCAACGAGAAATGCTGATGAGTGA 300
 Db 241 GGAAGCATGTCTCATGTGCAAGCACTCTAGCTCAACGAGAAATGCTGATGAGTGA 300

QY 301 AAATCGSAGCCAAAGAGGCTCATGGCTTTGAGGAGAGAGATCGCGATCGTCAG 360
 Db 301 AAATCGSAGCCAAAGAGGCTCATGGCTTTGAGGAGAGAGATCGCGATCGTCAG 360
 QY 361 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 420
 Db 361 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 420
 QY 421 AAAGAAATCGGCTTACTCATGATGTTTAAACAGAGCAATGGTTTTGACACACCGGTTT 480
 Db 421 AAAGAAATCGGCTTACTCATGATGTTTAAACAGAGCAATGGTTTTGACACACCGGTTT 480
 QY 481 TACTGGGAAACGCTTGGCGGTATCGCTTGAACCAATCTCTTATAGTACAAAACAAATCAGA 540
 Db 481 TACTGGGAAACGCTTGGCGGTATCGCTTGAACCAATCTCTTATAGTACAAAACAAATCAGA 540
 QY 541 AGAGCCGAGACGGAACGAGATGATCTCAAGCCAGATGGTTGAGTTCTTGAAGAAACAG 600
 Db 541 AGAGCCGAGACGGAACGAGATGATCTCAAGCCAGATGGTTGAGTTCTTGAAGAAACAG 600
 QY 601 AGTAGTAACGAACCTGTGTTTGTTCGTGAGTTGCTTAAACCGCGCTCGCTTAAACACATG 660
 Db 601 AGTAGTAACGAACCTGTGTTTGTTCGTGAGTTGCTTAAACCGCGCTCGCTTAAACACATG 660
 QY 661 ATGTGCTCTGATATTCGGAACAAGATAGAGCTTGAAGAAACCAATGTTGAGTTAGCTGA 720
 Db 661 ATGTGCTCTGATATTCGGAACAAGATAGAGCTTGAAGAAACCAATGTTGAGTTAGCTGA 720
 QY 721 ATGTCGAGAAAGATTATGATTTGCTCGGAACGTTGAAATGGAATGACATGACCTTCTTG 780
 Db 721 ATGTCGAGAAAGATTATGATTTGCTCGGAACGTTGAAATGGAATGACATGACCTTCTTG 780
 QY 781 CTATCGAGTTTATCTCTCAAGAATCTCGGTCTAGATGTTTCCACATCTGATCCAAAGGTA 840
 Db 781 CTATCGAGTTTATCTCTCAAGAATCTCGGTCTAGATGTTTCCACATCTGATCCAAAGGTA 840
 QY 841 AACCGGTTTATCCCGGATTAATCCGAACACCGTATCAACCGGTGATTTGCTCGT 900
 Db 841 AACCGGTTTATCCCGGATTAATCCGAACACCGTATCAACCGGTGATTTGCTCGT 900
 QY 901 GATTTGCTGACGTTTTGCTCTCCCTCCATGATGTTCAATTAATTCGACCCGAGATA 960
 Db 901 GATTTGCTGACGTTTTGCTCTCCCTCCATGATGTTCAATTAATTCGACCCGAGATA 960
 QY 961 ATGCGCGTTCTTTGGGAGATGATTTAGAGGAAACAGACACATGTCGGCTTAAATCGAG 1020
 Db 961 ATGCGCGTTCTTTGGGAGATGATTTAGAGGAAACAGACACATGTCGGCTTAAATCGAG 1020
 QY 1021 TGAATCTCTGCTAGATGATCTTCAATCCAGATATGCAATCAACGGTAAACAAACGAGCTG 1080
 Db 1021 TGAATCTCTGCTAGATGATCTTCAATCCAGATATGCAATCAACGGTAAACAAACGAGCTG 1080
 QY 1081 GATCAATAGTATCGGGAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATAT 1140
 Db 1081 GATCAATAGTATCGGGAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATAT 1140
 QY 1141 CTAAACGCTGTGTAAGAAAGATATGAGGCTTCACTCTCCAGAGCCCACTTCTATCATAG 1200
 Db 1141 CTAAACGCTGTGTAAGAAAGATATGAGGCTTCACTCTCCAGAGCCCACTTCTATCATAG 1200
 QY 1201 GCCCGTTTGGCCATTAACAGACAGATGTTGATGATGTTGTTTCCGAGAGGAGACACA 1260
 Db 1201 GCCCGTTTGGCCATTAACAGACAGATGTTGATGATGTTGTTTCCGAGAGGAGACACA 1260
 QY 1261 GCAATGATGAACATGATGGGCGGTAATGATGATGATGATGATGATGATGATGATGATG 1320
 Db 1261 GCAATGATGAACATGATGGGCGGTAATGATGATGATGATGATGATGATGATGATGATG 1320
 QY 1321 TTTTAACTTGAAGAGTTTGTGGAAGAAAGATGAGGTTGATGATGATGATGATGATGATG 1380
 Db 1321 TTTTAACTTGAAGAGTTTGTGGAAGAAAGATGAGGTTGATGATGATGATGATGATGATG 1380

QY 1381 GATTGAGACCTTGACCTTTTCGGGTCGGTCGTGCGATTTGCCCCGGGAGAAATCTTGGT 1440
 Db 1381 GATTGAGACCTTGACCTTTTCGGGTCGGTCGTGCGATTTGCCCCGGGAGAAATCTTGGT 1440
 QY 1441 TTTTACTACCTTATGTTTGAACGCGCATGATGTTTGAATGATGATGATGATGATGATGATG 1500
 Db 1441 TTTTACTACCTTATGTTTGAACGCGCATGATGTTTGAATGATGATGATGATGATGATGATG 1500
 QY 1501 GATGATAACGCGGTGACTTATCTGAGAACTGAGAGCTTTCTTTCGAGATGATGATGATG 1560
 Db 1501 GATGATAACGCGGTGACTTATCTGAGAACTGAGAGCTTTCTTTCGAGATGATGATGATG 1560
 QY 1561 CTTCCTGCTAATTTGCGCGCTGAGGCGCATTTAA 1593
 Db 1561 CTTCCTGCTAATTTGCGCGCTGAGGCGCATTTAA 1593

RESULT 3
 US-10-022-025a-1
 ; Sequence 1, Application US/10022025A
 ; Publication No. US2003092014A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Coleman, John R.
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Polypeptides for Catabolism of Absci
 ; TITLE OF INVENTION: Acid
 ; FILE REFERENCE: 3310 0003
 ; CURRENT APPLICATION NUMBER: US/10/022,025A
 ; PRIOR FILING DATE: 2003-01-09
 ; PRIOR FILING DATE: 2000-12-13
 ; NUMBER OF SEQ ID NOS: 8
 ; SOFTWARE: PatentIn version 3.0
 ; SEQ ID NO 1
 ; LENGTH: 2009
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; NAME/KEY: gene
 ; LOCATION: (1)..(2009)
 US-10-022-025a-1

Query Match 85.4%; Score 1360; DB 15; Length 2009;
 Best Local Similarity 87.7%; Pred. No. 0;
 Matches 1593; Conservative 0; Mismatches 0; Indels 223; Gaps 1;

QY 1 ATGGCTAGAAATCGAAAGCTCCTTAATCTTTCGCTTTCGCAATGAGAGGTTCTA 60
 Db 1 ATGGCTAGAAATCGAAAGCTCCTTAATCTTTCGCTTTCGCAATGAGAGGTTCTA 60
 QY 61 AGCCAAACCACTTGCCTTCTCCCTCTGCGGTCAATCATCTGAGTCCCATATCT 120
 Db 61 AGCCAAACCACTTGCCTTCTCCCTCTGCGGTCAATCATCTGAGTCCCATATCT 120
 QY 121 CTCTCTTATGACCTAATCCCGGTGACCGCTGAGGAGAAATCTCTTCGCGCGGTTA 180
 Db 121 CTCTCTTATGACCTAATCCCGGTGACCGCTGAGGAGAAATCTCTTCGCGCGGTTA 180
 QY 281 CTCTCTTATGACCTAATCCCGGTGACCGCTGAGGAGAAATCTCTTCGCGCGGTTA 340
 Db 281 CTCTCTTATGACCTAATCCCGGTGACCGCTGAGGAGAAATCTCTTCGCGCGGTTA 340
 QY 181 ATATCCGCTTATCAAAAACCGGAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 240
 Db 181 ATATCCGCTTATCAAAAACCGGAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 240
 QY 341 ATATCCGCTTATCAAAAACCGGAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 400
 Db 341 ATATCCGCTTATCAAAAACCGGAAGTTATCCCGGTCCAAAAGCTTCCCTTGGT 400
 QY 241 GGAAGCATGATCATCATATCTCAAGCATCTTACTCAACGAGCAATCCGTGATGACGTGAG 300
 Db 241 GGAAGCATGATCATCATATCTCAAGCATCTTACTCAACGAGCAATCCGTGATGACGTGAG 300
 QY 401 GGAAGCATGATCATCATATCTCAAGCATCTTACTCAACGAGCAATCCGTGATGACGTGAG 460
 Db 401 GGAAGCATGATCATCATATCTCAAGCATCTTACTCAACGAGCAATCCGTGATGACGTGAG 460
 QY 301 AAATCGSAGCCAAAGAGGCTCATGGCTTTGAGGAGAGATCGCGATCGTCAG 360
 Db 301 AAATCGSAGCCAAAGAGGCTCATGGCTTTGAGGAGAGATCGCGATCGTCAG 360
 QY 461 AAATCGSAGCCAAAGAGGCTCATGGCTTTGAGGAGAGATCGCGATCGTCAG 520
 Db 461 AAATCGSAGCCAAAGAGGCTCATGGCTTTGAGGAGAGATCGCGATCGTCAG 520
 QY 361 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 420
 Db 361 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 420
 QY 420 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 480
 Db 420 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 480
 QY 480 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 540
 Db 480 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 540
 QY 540 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 600
 Db 540 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 600
 QY 600 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 660
 Db 600 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 660
 QY 660 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 720
 Db 660 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 720
 QY 720 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 780
 Db 720 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 780
 QY 780 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 840
 Db 780 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 840
 QY 840 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 900
 Db 840 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 900
 QY 900 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 960
 Db 900 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 960
 QY 960 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1020
 Db 960 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1020
 QY 1020 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1080
 Db 1020 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1080
 QY 1080 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1140
 Db 1080 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1140
 QY 1140 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1200
 Db 1140 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1200
 QY 1200 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1260
 Db 1200 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1260
 QY 1260 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1320
 Db 1260 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1320
 QY 1320 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1380
 Db 1320 TGCATATCCGACGTAGGAGAAAGATTTCTGAATAGCCCGGTTTTGCTGATGACCGGTT 1380

QY 421 AAGAAATCGGCTTACTCACTGATGTTTAAACAGAGCAATTGCTTTTGCACACACGGTGT 480
 DB 581 AAGAAATCGGCTTACTCACTGATGTTTAAACAGAGCAATTGCTTTTGCACACACGGTGT 640
 QY 481 TACTGGGAAACGCTTGGCGGTATCGCTTGAACCACTCTTTAGTAAACAAATCAATGGA 540
 DB 641 TACTGGGAAACGCTTGGCGGTATCGCTTGAACCACTCTTTAGTAAACAAATCAATGGA 700
 QY 541 AGAGCCGAGACGCGACGACGAGTATCTCAACCCAGATGCTTGAATTTCTTGAAAAACAG 600
 DB 701 AGAGCCGAGACGCGACGACGAGTATCTCAACCCAGATGCTTGAATTTCTTGAAAAACAG 760
 QY 601 AGTAGTAAAGAACCCGCTTTGTGCTGAGTGTGCTTAAACGGCGCTGCTTAAACAGATG 660
 DB 761 AGTAGTAAAGAACCCGCTTTGTGCTGAGTGTGCTTAAACGGCGCTGCTTAAACAGATG 820
 QY 661 ATGTGCTCTGTATTCGACAAAGATATGAGCTTGAACAAAAACATGTTGAGTTAAGTAA 720
 DB 821 ATGTGCTCTGTATTCGACAAAGATATGAGCTTGAACAAAAACATGTTGAGTTAAGTAA 880
 QY 721 ATGTGCTCTGTATTCGACAAAGATATGAGCTTGAACAAAAACATGTTGAGTTAAGTAA 780
 DB 881 ATGTGCTCTGTATTCGACAAAGATATGAGCTTGAACAAAAACATGTTGAGTTAAGTAA 940
 QY 781 CTATCGAGTGTGATCTCTAAAGACTCCGGTCTAGATGTTTCAACTGCAACCAAGTAA 840
 DB 941 CTATCGAGTGTGATCTCTAAAGACTCCGGTCTAGATGTTTCAACTGCAACCAAGTAA 1000
 QY 841 AACCGGTTGATCCCGGATTTATCCGAAACCGGATTTCAACCGGATTTGCTGCTGCT 900
 DB 1001 AACCGGTTGATCCCGGATTTATCCGAAACCGGATTTCAACCGGATTTGCTGCTGCT 1060
 QY 901 GATTGCTGACGCTTTGCTCTCCCTCCATGCTTCAATTAATTAATTCGACCCGAGACATA 960
 DB 1061 GATTGCTGACGCTTTGCTCTCCCTCCATGCTTCAATTAATTAATTCGACCCGAGACATA 1120
 QY 961 ATGCGCGTCTTTG----- 974
 DB 1121 ATGCGCGTCTTTGCGGTATGACACACATTATTGATTAATTAATTAATTTGCT 1180
 QY 975 ----- 974
 DB 1181 TGAATAATGCTTAGATTAATTAATTAACATGAAATTTGAGATGACGCTGT 1240
 QY 975 ----- 974
 DB 1241 TGTGAATAATTAAGCATTAGAATTTTGTGACATCAAAATTAGTAATTTAGAT 1300
 QY 975 -----GGA 977
 DB 1301 TTTAATAAGCTTTCTAATAAAAGTAAAGTCTAATAATTTTGAATTAATTAACAGA 1360
 QY 978 GATGATATTCAGAGAAACAGACAGTGGCTTTAATCGAGTATCTCGCTAGAGAT 1037
 DB 1361 GATGATATTCAGAGAAACAGACAGTGGCTTTAATCGAGTATCTCGCTAGAGAT 1420
 QY 1038 GGTCTCTTATCCAGATATGCAATCAACGCTTCAAAACGATGCTCAAGTATGCGGAA 1097
 DB 1421 GGTCTCTTATCCAGATATGCAATCAACGCTTCAAAACGATGCTCAAGTATGCGGAA 1480
 QY 1098 ATCAAGAGCCTAGATGAACTGACTGCTTCACTTCAATCAATCAACGCTGCTGAA 1157
 DB 1481 ATCAAGAGCCTAGATGAACTGACTGCTTCACTTCAATCAATCAACGCTGCTGAA 1540
 QY 1158 AGAAGTATGAGGCTTCACTTCCAGGCTTCTTATCAAGGCTGCTTGGCTTAAAC 1217
 DB 1541 AGAAGTATGAGGCTTCACTTCCAGGCTTCTTATCAAGGCTGCTTGGCTTAAAC 1600
 QY 1218 AGACAGATGCTTATGCTGCTTGTTCGCGACGAGACACAGCAATGCTGAACATG 1277
 DB 1601 AGACAGATGCTTATGCTGCTTGTTCGCGACGAGACACAGCAATGCTGAACATG 1660
 QY 1278 GCGCGATGCGATGATCAACGCTGCGGTATGCTTTGAGTTAAACCTGAGAGTT 1337

DB 1661 GCGCGATGCGATGATCCACACGCTGCGGTATGCTTTGAGTTAAACCTGAGAGTT 1720
 QY 1338 CGTGCAAAAGAGATGAGTGAAGTGTGAGTTTGGCTCTTGGGCGATTTGAGATGAC 1397
 DB 1721 CGTGCAAAAGAGATGAGTGAAGTGTGAGTTTGGCTCTTGGGCGATTTGAGATGAC 1780
 QY 1398 TTTGCGGTGCGGTCTGCGATTTTGGCCCGGAGAAATCTTGTGTTTACTACCGTTATGT 1457
 DB 1781 TTTGCGGTGCGGTCTGCGATTTTGGCCCGGAGAAATCTTGTGTTTACTACCGTTATGT 1840
 QY 1458 TTGGAACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1517
 DB 1841 TTGGAACGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1900
 QY 1518 CTATCTGAGAAACTGAGGCTTTCTTGGAGATGCTTATCTTCTGCTTAAATGCG 1577
 DB 1901 CTATCTGAGAAACTGAGGCTTTCTTGGAGATGCTTATCTTCTGCTTAAATGCG 1960
 QY 1578 CCGTAGGCGCACTTAA 1593
 DB 1961 CCGTAGGCGCACTTAA 1976

RESULT 4
 US-09-349-385-2
 ; Sequence 2, Application US/09349385
 ; Patent No. US20020152495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ito, Toshio
 ; APPLICANT: Fromm, Michael
 ; APPLICANT: Meyerowitz, Elliot
 ; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
 ; FILE REFERENCE: MBI-0002
 ; CURRENT APPLICATION NUMBER: US/09/349,385
 ; EARLIER FILING DATE: 1999-07-09
 ; EARLIER APPLICATION NUMBER: 60/115,967
 ; EARLIER FILING DATE: 1999-01-15
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: Patenceln Ver. 2.0
 ; SEQ ID NO 2
 ; LENGTH: 1902
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; FEATURE:
 ; OTHER INFORMATION: cDNA
 ; US-09-349-385-2

Query Match 61.3%; Score 976; DB 10; Length 1902;
 Best Local Similarity 78.0%; Pred. No. 0;
 Matches 1238; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCTTAATCTTTGACCCCTTTGTCAAATGACAGCTTAAAGCAACCAACCTTGC 78
 DB 90 AATTAATCTTTGACCCCTTTGTCAAATGACAGCTTAAAGCAACCAACCTTGC 129
 QY 79 TTTCTCTCTCTGCGCGTCAATCATCTGCTGCGCAATCTCTCTTGAATGACCTAT 138
 DB 150 CTCTCTCTCTCTGAGCTCTCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 209
 QY 139 CCGGTGACGCTGCTTGGGAGAAATACCTTTCGCGGCTTAAATATCGGTTATCAAA 198
 DB 210 CCGGTGACGCTGCTTGGGAGAAATACCTTTCGCGGCTTAAATATCGGTTATCAAA 254
 QY 199 ACCGAAACGTTATTCCTGCTCAAAAGCTTCTCTTGTGGAACATGTCATG 258
 DB 255 CAAACCAACGTTATTCCTGCTCAAAAGCTTCTCTTGTGGAACATGTCATG 314
 QY 259 TCAGACATCTAGCTCACCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 318
 DB 315 TCAGACATCTAGCTCACCGAGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 374
 QY 319 CTATGCTTCAAGCTTGAAGAGACTGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 378

```
Db 375 TTAATGGGCGTTAGTTGGAGAACTGCGATGCTGCAAGTGAATCTTGATGACT 434
Qy 379 AAAGAGATTCGAATAGCCCGGTTTTTCTGATGCAACGGGTTAAAGAAATCGCTTACTCA 438
Db 435 AAAGAGATTCGAATAGCCCGGTTTTTCTGATGCAACGGGTTAAAGAAATCGCTTACTCC 494
Qy 439 CTGATGTTTAAACAGAGCAATGGTTTGGACCAACAGGTTTACTTGCGGAAACGGCTTGC 498
Db 495 CTGATGTTTAAACAGAGCAATGGTTTGGACCAACAGGTTTACTTGCGGAAACGGCTTGC 554
Qy 499 CGTATCGCTTCAACCATCTCTTATGTAACAAACAAATCAGAGAGCGGAGCGCAAGA 558
Db 555 AAAATGCGCTTATATCTTTTCAAGCCGAAACAGATTAACGTTCCGAAACGCAAGAGA 614
Qy 559 CGATGATCTCAAGCCAGATGTTGATGTTCTTGAACAAACAGATGTAAGCA---CCG 615
Db 615 ACGGTATGCGCAATCAAAATCGTAAGTGTCTCAACAAACAGATGTAAGCAACCAAGGCTC 674
Qy 616 TGTGTTGCTGATGCTTAAACCGCGCTTAAACGATGATGCTGCTGATTC 675
Db 675 TGTGTTGCTGATGCTTAAACCGCGCTTAAACGATGATGCTGCTGATTC 734
Qy 676 GGAAGAGATGAGCTTGAACAAACCAATG-----TTGAGTTAGTGAATGCTC 726
Db 735 GGAAGAGATGAGCTTGAACAAACCAATG-----TTGAGTTAGTGAATGCTC 794
Qy 727 GGAAGAGATGATGCTGCGGAGCTTGAATGCTGATGATGCTGCTGCTGCTGCTGCTGCTG 786
Db 795 GGAAGAGATGATGCTGCGGAGCTTGAATGCTGATGATGCTGCTGCTGCTGCTGCTGCTG 854
Qy 787 GAGTTGATCTCAAAAGCTCGGCTGATGTTCAACTGCTGACAAAGTTAAACCGG 846
Db 855 GAATTTGATCTCAAAAGCTCGGCTGATGTTCAACTGCTGACAAAGTTAAACCGG 914
Qy 847 TTTGTTATCCCGGATGATGCTGCGGAGCTTGAATGCTGATGCTGCTGCTGCTGCTGCTGCTG 906
Db 915 TTTGTTATCCCGGATGATGCTGCGGAGCTTGAATGCTGATGCTGCTGCTGCTGCTGCTGCTG 974
Qy 907 GTGAGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 966
Db 975 GTTGAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1034
Qy 967 GTTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1026
Db 1035 GTTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
Qy 1027 CTGCTGATGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1086
Db 1095 CTGCTGATGATGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
Qy 1087 GTATCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1146
Db 1155 ATCTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1214
Qy 1147 GCTGTTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1206
Db 1215 GCTGTTGGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1274
Qy 1207 TTGCGCATTAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266
Db 1275 TTGCGCATTAACAGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334
Qy 1267 GTGAACATGTTGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
Db 1335 GTGAACATGTTGGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
Qy 1327 CCTGAGAGGTTGCTGCGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1386
Db 1395 CCTGAGAGGTTGCTGCGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1454
Qy 1387 AGACTTGACCTTTCGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1446
```

```
Db 1455 AGGCTTGACCGCTTGGGCTCGGCTCGGCTTGGCCCGGAGAAATCTTGCTTGACC 1514
Qy 1447 ACCGTATGTTTGAACGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1503
Db 1515 ACCGTATGTTTGAACGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1574
Qy 1504 GGTAAAGCGCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1563
Db 1575 GAGAAGACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1634
Qy 1564 CTGCTTAATTTGGCGGTAGCGCAGTT 1591
Db 1635 GCTGCTTAATTTAGCCCGGAGCCAGTT 1662

RESULT 5
US-09-349-385-3
; Sequence 3, Application US/09349385
; Patent No. US20020152495A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Toshiro
; APPLICANT: Fromm, Michael
; APPLICANT: Meyerowitz, Elliot
; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
; FILE REFERENCE: MBI-0002
; CURRENT FILING DATE: 1999-07-09
; EARLIER FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent Ver. 2.0
; SEQ ID NO 3
; LENGTH: 7869
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: genomic
US-09-349-385-3

Query Match 49.8%; Score 794; DB 10; Length 7869;
Beet Local Similarity 70.3%; Pred. No. 7,66-259;
Matches 1238; Conservative 0; Mismatches 320; Indels 202; Gaps 5;
```

```
Qy 19 AGCTCTTAATCTTTGCTTGTGCAATGACGCTTCTTAAACCAACCACTTGGC 78
Db 2154 AGTACCTTACTTTGGCCCTCTTGCAATGATGATGATGATGATGATGATGATGATGATGATG 2213
Qy 79 TTCTCCCTCCCTGCGGCTGACATCATGCTGCTGCGCATGATCTCTTCTTATGACCTAT 138
Db 2214 CTCTCTCTCTGATGCTCTCTGCTGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2273
Qy 139 CCGGTGACCTGCTTGGGAGAAATACCTCTTGGCGGCTTAATATCCGTTACTACAAA 198
Db 2274 CCGGAGAGACCGGATGAGGAAAGTACTTCTCCACCG-----CCGCGCT 2318
Qy 199 ACCGAAACGTTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 258
Db 2319 CAACACACCGGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2378
Qy 259 TCAGACATCTAGCTCACCGACGATCGCTGATGACGCTGAGAAATTCGAGCCAGAGG 318
Db 2375 TCAGACATCTAGCTCACCGACGATCGCTGATGACGCTGAGAAATTCGAGCCAGAGG 2438
Qy 319 CTATAGCTTTAGCTTGAAGAGACTGCGGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 378
Db 2439 TTAATGCGTTTATGTTGGAGAACTGCGGATGCTGATGCTGATGCTGATGCTGATGCTGATG 2498
Qy 379 AAAGAGATTCGAATAGCCCGGTTTTTCTGATGCAACGGGTTAAAGATTCGGCTTACTCA 438
Db 2499 AAAGAGATTCGAATAGCCCGGTTTTTCTGATGCAACGGGTTAAAGATTCAGCTTATTC 2558
Qy 439 CTGATGTTTAAACAGAGCAATGCTTGTGCAACAGCTGTTTACTTGCGGAAACGCTTGGC 498
```

Db 2559 CTCATGTTAAACGTCATACGCTTCCGCTTACGCGCTTACTGCGAACCCTTGAGA 2618
 QY 499 CGATCGCTTCGAAACCATCTCTTAACTAGTACAAAACAAATCAGAGAGCCGAGCAACA 558
 Db 2619 AAATCGCGCTTAACTATCTTTTCAAGCCGAAACGATTAACGTTCCGAAACGAGAGA 2678
 QY 559 CGATGATCTCAGCCAGATGCTGAGTTCTTGAAGAAACAGATGATGAGAA---CCC 615
 Db 2679 AGCGTATCGCGAATCAAAATCGGAAGTCTCAAAAACAGATGAACCAAGGCTC 2738
 QY 616 TGTGTTGTCGATGCTTAAACCGCGCTGCTTAACACATGATGCTGATTC 675
 Db 2739 TGTTCGACGATGATGATCAAAACGCGATCGCTTAATACATGATGCTGATTC 2798
 QY 676 GGAAGAGATGATGCTGAAAAACATG-----TGAATTCGTAATGATC 726
 Db 2799 GAAAAAGATACGAGCTGAGAGAGATGAAGAAAGATGATGATGATGATGATG 2858
 QY 727 GAAAGATGATGATGCTGAGAACGCTGAATGATGATGATGATGATGATGATG 786
 Db 2859 GAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2918
 QY 787 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 846
 Db 2919 GAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2978
 QY 847 TTTGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 906
 Db 2979 TTTGATCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3038
 QY 907 GTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 966
 Db 3039 GTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3098
 QY 967 GTTCTTTG----- 974
 Db 3099 GTTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3158
 QY 975 ----- 974
 Db 3159 AAAAATTCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 3218
 QY 975 ----- 974
 Db 3219 AATTTAATATGATTTAATTAATTAATTAATTAATTTGATTTTATTAACAACA 3278
 QY 975 GAGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3338
 Db 3279 GAGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3398
 QY 1035 GATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1094
 Db 3339 GATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3398
 QY 1095 GATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1154
 Db 3399 GATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3458
 QY 1155 GATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1214
 Db 3459 GATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3518
 QY 1215 AATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1274
 Db 3519 AATGATGATTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3578
 QY 1275 GTGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334
 Db 3579 GTGAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3638
 QY 1335 GTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
 Db 3639 GTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3698

QY 1395 ACCCTTCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1454
 Db 3699 ACCCTTCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3758
 QY 1455 GTTTCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
 Db 3759 GTTTCGAGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3818
 QY 1512 CGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1571
 Db 3819 CGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3878
 QY 1572 ATTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1591
 Db 3879 ATTGCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3898
 RESULT 6
 US-09-349-385-7
 ; Sequence 7, Application us/09349385
 ; Patent No. US20020152495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ico, Toshio
 ; APPLICANT: Fromm, Michael
 ; APPLICANT: Meyerowitz, Elliot
 ; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
 ; FILE REFERENCE: MBI-0002
 ; CURRENT APPLICATION NUMBER: US/09/349,385
 ; EARLIER APPLICATION NUMBER: 60/115,967
 ; EARLIER FILING DATE: 1999-07-09
 ; NUMBER OF SEQ ID NOS: 13
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 7
 ; LENGTH: 1611
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: cDNA
 US-09-349-385-7
 Query Match 40.9%; Score 651.6; DB 10; Length 1611;
 Best Local Similarity 66.5%; Pred. No. 1e-210;
 Matches 1025; Conservative 0; Mismatches 484; Indels 33; Gaps 5;

QY 481 TACTGCGAAACGTTGCGCGTATGCTTGAAACATCTCTTTAGTACAAACAAATCAGA 540
 Db 473 TACTGCGAAAGCTTCAGGAAATCGCTCTAAATCACTCTTTCGCCCCGCGCAAGATAAA 532
 QY 541 AGAGCCGACGACGACGAGATGCTCAAGCCAGATGATGATTTCTTGAACAAAACAG 600
 Db 533 GCGCTGAGCTCCAAAGCTCTCAATTCGCGCCCAATAGTTTACATCTCT---AAATAAC 589
 QY 601 AGTAGTACGAAACCTGTTTGTCTGAGTTGCTTAAACGCGCTCGCTTACCAACATG 660
 Db 590 AAGCGCCACCGCAGCTTACGTTTCCCAAGTCTTAAAGGCTTCGCTCAGTAAACAG 649
 QY 661 ATGTGCTGTATTCGGAACAAGATGATGAGCTTGAACAAAACCA-----TGTAG 711
 Db 650 ATGTGCTGTATTCGGAACAAGATGATGAGCTTGAACAAAACGCGAATGGAAC 709
 QY 712 TTACGTAAATGCTGAAAGATGATGATGCTGGAACGTTGATGATGATGATGATGAT 771
 Db 710 CTGGAATATGATGACCAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
 QY 772 CTTCCTGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 831
 Db 770 CTTCCTGCTATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
 QY 832 CCAAGGTAACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 891
 Db 830 CCAAGGTAACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 889
 QY 892 TTGCTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 951
 Db 890 ACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 949
 QY 952 CCGGACATATCCCGGTTCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1011
 Db 950 TCGGACATATCCCGGTTCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGAT 1009
 QY 1012 TTATGAGTGAATCTCTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1071
 Db 1010 TTATGAGTGAATCTCTGCTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 1069
 QY 1072 AAGAGTGAATCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131
 Db 1070 GAGAGTGAATCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1129
 QY 1132 CTTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1191
 Db 1130 ATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1189
 QY 1192 CTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1251
 Db 1190 CTATCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1249
 QY 1252 GGGACCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1311
 Db 1250 GGGACCAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1309
 QY 1312 CTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1371
 Db 1310 CCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1369
 QY 1372 CTTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1431
 Db 1370 CTTGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1429
 QY 1432 AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1491
 Db 1430 AATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1489
 QY 1492 GAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1551
 Db 1490 GAGAGTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1549

QY 1552 GCTAATCTCTTCTGCTAAATTCGCGCGTAGGCGAGTTAA 1593
 Db 1550 GCTAATCTCTTCTGCTAAATTCGCGCGTAGGCGAGTTAA 1591

RESULT 7
 US-10-163-198-20
 ; Sequence 20, Application US/10163198
 ; Publication No US20030126645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebecca E. Cahoon
 ; APPLICANT: Elmer P. Heppard
 ; APPLICANT: No. US20030126645A1uhiro Nagasawa
 ; TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
 ; FILE REFERENCE: B01487 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,198
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 60/295,921
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 60/334,317
 ; PRIOR FILING DATE: 2001-11-28
 ; NUMBER OF SEQ ID NOS: 111
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 20
 ; LENGTH: 1764
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 US-10-163-198-20

Query Match 37.6%; Score 599.6; DB 15; Length 1764;
 Best Local Similarity 65.0%; Pred. No. 6.2e-193; Indels 42; Gaps 7;
 Matches 1014; Conservative 0; Mismatches 504;

QY 64 CAACCAACCTTCT 120
 Db 133 CAAGGAACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 192
 QY 121 CTCTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 174
 Db 193 TCTCTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 252
 QY 175 CGGTATATCCGCTTCT 222
 Db 253 TGAAGAAACCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 312
 QY 223 AAAGGCTTCCCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 282
 Db 313 AGAGGCTTCCCTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 369
 QY 283 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
 Db 370 ATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429
 QY 343 ACTGCGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 402
 Db 430 ACAAGCGCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 489
 QY 403 TTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
 Db 490 TTTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 549
 QY 463 TTTGACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 522
 Db 550 TTTGCGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 609
 QY 523 AGTACAAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 582
 Db 610 TGGCCCAACAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
 QY 583 GAGTTCTTGAACAAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 642
 Db 670 AACTCATTCGAAATCAACGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 729

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.inpb

Page 8

QY 643 GCCTGCTTACACATGATGCTCTGTTTCGACAGAGATGAGCTTGAACAAAC 702
Db 730 GCGTCACTGAAACATGATGCTGCTGTTTGAACAAAGTACAACTTGACAGATA 789
QY 703 CA-----TGTGAGTTACGTGAATGCTGAAGAAGTTATGCTCGGACG 753
Db 790 AACACGCAATGACGAGCTATCCATGTTGGTGAACAAAGCTATGACTGTTGGGACCC 849
QY 754 TTGAATTGACATGATCCTCTTGGCTATGAGAGTTGATCTCAAAAGCTCGGCT 813
Db 850 CTTAATTTGGGAGAACATATCTCTTTCGAAAGAGTTGACCTACAGAAATCCGGTTC 909
QY 814 AGATGTTCCACTGCTACCAAGGTAACCGGTTTGTATCCGAGTATATCCGACAC 873
Db 910 ACCGTCTCCAAATTAATGCTCCTCAAGTGAACCGGTTGTTGTTCAATCATGCGACAC 969
QY 874 CGTATGAAACCGGCTGATTTGCTGATTTGCTGACAGTTTCTCTCCCTCATGCT 933
Db 970 CAGGCGACCAACCAACCAACCGGATTTGCTCATGTTTCTCTCTCCAAAGT 1029
QY 934 TCAGATTAATTCGACCCGACATATCGCCGTTCTTGGAGATGATATTCGAGGA 993
Db 1030 CCGATTAATTTGCTCACTGCTGACATGATTTGCTGCTGCTGGAATGATTTTGAAGGG 1089
QY 994 ACAGACACAGTTGCTGTTTATTCAGTGAATCTGCTGATGATGCTCTTCTCATCAGAT 1053
Db 1090 ACCGACACGCTGCTGTTTGAATGATGATGATGCTGCGAGATGCTCTCATCCGAG 1149
QY 1054 ATGCAATCAACGCTGACAAACGAGCTGATCAAGTATGCTGGAATTCAGAGCCCTAGAT 1113
Db 1150 GTGCAAGGAGATGACAGAGAGTTGACGCGGTGTTAGG---GTTGGGCTTTGACG 1206
QY 1114 GAATTCGATCTGCTTCACTCATCTTAAACGCTGCTGCTGATGAAAGATTTAGAGCTT 1173
Db 1207 GAGAGGTCGTGCGGACGAGGCTGATCTTGGCGGCTGATGAAAGAGCTTCTGAGCTG 1286
QY 1174 CATCTTCAGGCGCCACTTATCATGAGGCGCTTGGCCATTAACAGACATGATGAT 1233
Db 1267 CACCGCGCGGCGCTTCTCTGCTGCGCGCTTGGCCATCATGATGACCATGAT 1326
QY 1234 GCTGCTGTTGCTGCGGACGAGGACCAAGATGCTGATGATGCTGCTGCTGATGAT 1293
Db 1327 GGGATTCACGCTGCTGCGGAGCAACGCTATGCTTAAATGCTGCGCAATGCAAGGAC 1386
QY 1294 CCAACGCTGCTGCTGATCTTGTGAGTTAAACCTGAGAGGTTGCTGCGCAAAAGAGT 1353
Db 1387 CCGAGGCTGCTGCTGACCACTTGAATTCAGGCGCGAGAGGTTATGAGTCTGAA--- 1443
QY 1354 GAGGTGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1413
Db 1444 ---AAGAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1500
QY 1414 CGGATTTGCGCGGAGAGATCTTGTGTTTACCTGATGCTGCTGCTGCTGCTGCTGCT 1473
Db 1501 AGAACAATGCGCGGAGAGATCTTGTGTTTGAACGCTGCTGCTGCTGCTGCTGCTGCT 1560
QY 1474 TTTCATGAGTTTGAAGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1533
Db 1561 TTGCAATGAGTTTGAAGGAGGAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1620
QY 1534 AGGCTTTCTTGCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1593
Db 1621 AGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1680

RESULT 8
US-10-163-198-24
; Sequence 24, Application US/10163198
; Publication No. US2003012645A1
; GENERAL INFORMATION:
; APPLICANT: Rebecca E. Cahoon
; APPLICANT: Elmer P. Heppard

APPLICANT: No. US2003012645A1uhio Nagasawa
APPLICANT: Hajime Sakai
TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
FILE REFERENCE: B01487 US NA
CURRENT APPLICATION NUMBER: US/10/163,198
PRIOR FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/295,921
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 60/334,317
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Microsoft Office 97
SEQ ID NO 24
LENGTH: 1905
TYPE: DNA
ORGANISM: Glycine max
US-10-163-198-24
Query Match 36.8%; Score 585.8; DB 15; Length 1905;
Best Local Similarity 66.1%; Pred. No. 3.4e-188;
Matches 934; Conservative 0; Mismatches 452; Indels 27; Gaps 5;
QY 193 TCAAAACCGGAAAGTATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 252
Db 294 TCCAAACACCTTAAATGATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 353
QY 253 CTCATGTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 312
Db 354 CTCATGTCAGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 410
QY 313 AAGAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 372
Db 411 ACCAGGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 470
QY 373 GTAGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 432
Db 471 GTAGGAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 530
QY 433 TACTCATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 492
Db 531 TACAGCTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 590
QY 493 CTTGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 552
Db 591 CTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 650
QY 553 CAACGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 612
Db 651 CAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 704
QY 613 CCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 672
Db 705 GCTTTCGAAATCCGAGGCTTCTCAAGAGAGCTGCTCAACCAAGATGCTGCTGCTGCT 764
QY 673 TTCGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 733
Db 765 TTTCGCAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 824
QY 724 GTGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 783
Db 825 GTGCAAGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 884
QY 784 TCGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 843
Db 885 AAAAGCTTGAACCTTCAAAAATCCGCTTTCGCTGCTGCTGCTGCTGCTGCTGCTGCT 944
QY 844 CCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 903
Db 945 CCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1004
QY 904 TTGCTGCAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 963
Db 1005 TTGCTGCAAGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1064

QY 964 GCCGTTCTTTGGAGATGATTTCAAGAGAACAGACACAGTTGCGGCTTTAATCCAGTGG 1023
 DB 1065 GCTGTCCTTGGGAAATGATATTTAGGGGACCGACACGGTGGGTTTGATGATGG 1124
 QY 1024 ATCCGCTAGATGATGCTTTCATCCAGATATGCAATCAACGCTACAAAACGAGTGGAT 1083
 DB 1125 ATATGCGAAGATGATGCTTTCATCCGAGGTAACAAGAGGGTGCAGAGGAGCTGAC 1184
 QY 1084 CAAGTACT---CGGAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCCATAT 1140
 DB 1185 GCGGTGTTGAGGTGTGTGGCCGCTTTGAAAGAGAGACGTTGCGGAGACGGGTAT 1244
 QY 1141 CTACGAGCTGTGGTGAAGAGATTTAGAGCTTTCATCTCCAGGCCCACTTATCATGG 1200
 DB 1245 CTTCGCGGTGTGGAGAGAGGTTCTGAGGCTGCACCTCCAGGCCGCTTCTGTGG 1304
 QY 1201 GCCCGTTGGCATTAACAGACAGATGTTGATGTCTCTTGTTCGGGAGGACCA 1260
 DB 1305 GCCCGTTGGCATTAACAGACAGATGTTGATGTCTCTTGTTCGGGAGGACCA 1364
 QY 1261 GCATGCTGAACATGTGGGCGGTATGCGATGATCACAAGTGTGGTGTATCTTTGGAG 1320
 DB 1365 GCCATGCTTAATATGTGGGCGGTATGCGATGATCACAAGTGTGGTGTATCTTTGGAG 1424
 QY 1321 TTTAACTGAGAGCTTGTGGCAAAAAGAGTGAAGTGTGGTGTCTTGGGCTG 1380
 DB 1425 TTCAAGCCCGAGAGTTCATG-----GCCCTGAGGCGGAGTTCCTGCTCGGCTG 1478
 QY 1381 GATTTGAGACTTGACCTTTTCGGGTGGGTCTGTGGATTTGCCCGGAGAGATCTTGGT 1440
 DB 1479 GATGTGAGGTGTGTCCATTCGGGTGGGTGTGAAGAACCTGCCGAGAGCTTGGGT 1538
 QY 1441 TTTACTACCTTATGTTTGGACCGCATGATGTTTACATGATTTGAATGGGACCGTCC 1500
 DB 1539 TTGAGCACCGTGACTTCTGGGTGGCGAGGCTTTTGACAGATTTGAATGGTACATCT 1598
 QY 1501 GATGTGAACGCGCTTGACTTATCTGAAGAACTGAGGCTTTCTCCGAGATGGCTAATCT 1560
 DB 1599 GATGAGGAGGAGTGTGATCTTAACGAGGTGCTGAGGCTCTGTGTGAATGGCTAACCCG 1658
 QY 1561 CTTCCTGCTAATGCGCGGCTAGCGCAGTTAA 1593
 DB 1659 CTCTATGTTAAAGTTCCGCCCTAGCGCTGAGATTA 1691

RESULT 9

US-09-349-385-13

Sequence 13, Application US/09349385

Patent No. US20020152495A1

GENERAL INFORMATION:

APPLICANT: Ito, Toshio

APPLICANT: Fromm, Michael

APPLICANT: Meyerowitz, Elliot

TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT

FILE REFERENCE: MBI-0002

CURRENT APPLICATION NUMBER: US/09/349,385

EARLIER FILING DATE: 1999-07-09

EARLIER APPLICATION NUMBER: 60/115,967

NUMBER OF SEQ ID NOS: 13

SOFTWARE: Patent In Ver. 2.0

SEQ ID NO 13

LENGTH: 1799

TYPE: DNA

ORGANISM: Phalaenopsis sp.

FEATURE:

OTHER INFORMATION: cDNA

FEATURE:

OTHER INFORMATION: Hybrid SM9108

US-09-349-385-13

Query Match

22.3%; Score 355; DB 10; Length 1799;

Best Local Similarity 58.0%; Pred. No. 1,9e-109;
 Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTGGAGCCAGAGAGCTTCATGCGCTTTCAGGTTAGAGAGACTGCGGATGCTGACGTCG 363
 DB 273 TTGAGCGCTTCCGCTTCATGAGCTTCTTCGGTGGGCTCACTGCTTCACTGCTTCAAGC 332
 QY 364 ATCCGAGCTAGAGAGAGATTCGAAATAGCCGCTTTTGTGATGCAACCGCTTAA 423
 DB 333 CACCGAGAAACCGAAAGAGATCTTCAAGCCCACTTCGCTGATGCGCCCATTA 392
 QY 424 GATGAGCTTACTGATGATGTTTAAACAGCAATTTGGTTCGACACCGGTTTAC 483
 DB 393 GATTCAGATGCAACTTCTGTTTATGCGCTTATGAGGTTTGGCCATTTGGGAGATTAC 452
 QY 484 TGGCGAAGCTTGGCGCTTATGCTTCGAAACCATCTTTAGTACAAAACAAATCAGAGA 543
 DB 453 TGGAGAAACCTGAGAGAGATTTCCGCAATATCTTTCACTGCGCGCGAGTTTCAATCG 512
 QY 544 GCCGAGACGCAACGAGATGATCTCAAGCAGATGTTGATGTTTCTTGAAGAAACA 601
 DB 513 TTGAGAAAGCAACGAGATGAGATTTGGCGAAGAAATGGTGGATATGAAGAGATG 572
 QY 602 -GTAGTACGAACCTGTTTGTCTGTGAGTTGCTTAAACGGCGCTTACAAACATG 660
 DB 573 GAGAGAAATGAGATGTTGAGAGAGAGAGATTTGCACTACGGGTCTTTGAATACATC 632
 QY 661 ATGTCCTGTATTTGGAACAAGATGAGCTTGAAGAAACATGTTGATGCTGCA 720
 DB 633 ATGTTGACTGTTTGGGAGAAAGTTGATTTGCAAGATGAGGCTTGAAGCTTGAAG 692
 QY 721 ATGTC---GAAGAGTTATGATTTGCTCGAAGCTTGAATGAGTCACTGATCACTTCT 777
 DB 693 TTGATCTTAAAGAGATGATGATTTGCTGAGATTTCACTGAGGCTGATCATTTGCT 752
 QY 778 TGGCTATGAGATTTGATCTTCAAGACTCGGCTGATGATTTCACTGATCACTGATCAAG 837
 DB 753 CTTTGGAGATGATTTGATTTGCAAGGTGAGAGAAATGCAAGACCTTGTGCTAAG 812
 QY 838 GTAAACCGTTGATTTCCGATTTATCCGAACTATATCCGAAACCGTATCAAC--- 885
 DB 813 GTCAATGATTTGTAAGAGATCATAGACGCAATTAAGAGAGACCAAGCGCTAGG 872
 QY 886 -----GATGATTCCTGATTTGATTTGCTGCAAGTTTCTCTCCCTCATGTTCAAT 939
 DB 873 ATGATGAGGTGAGAGAGATTTGATTTGATTTGCTTGTGAGAGAGAAAT 932
 QY 940 AATTAATCCAGCCGCAATATGCGCTTCTTTGGAGAGATATTAAGAGAGACAG 999
 DB 933 AGACTCTCAAGATTCGATATGCTGCAAGTTCTTTGGAGAAATATCTTTAGAGAACTGAT 992
 QY 1000 ACAGTTGCGGTCTTATGATGATGATTCCTGCTAGAGATGTCCTTATCCAGATATGCA 1059
 DB 993 ACTGTTCCATCTTATGATGATGATGATGCTTGAAGTCTTCACTCATCTGATATCA 1052
 QY 1060 TCAACGTTACAAACAGCTGATCAAGTGTGCGAAATCAAGAGCCCTGATGATGATCT 1119
 DB 1053 TCGAAGGACAAAGTTGAGATGATTTCTGCTGATGCTTCAAGGCGCATTTGATGATCT 1112
 QY 1120 GACTGCTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1179
 DB 1113 GATATCAACGATCTTCTTATCTCACTTCACTTCACTTCACTTCACTTCACTTCACT 1172
 QY 1180 CCAAGCCCACTTCTATCATGATGAGGCGGTGCGCAATACAGACAGATGATGATGCT 1239
 DB 1173 CCGGAGCTCTATGATGATGAGGCTGCTGAGCTATTCAGATGATGATGATGCTCAC 1232
 QY 1240 CTGTTCCGCGAGGAGCCACAGATGAGAGATGAGGCGGCTATGATGATGATGATGATG 1299
 DB 1233 ATGATTCCTGCTGAGGAGCTGATGATGATGATGATGATGATGATGATGATGATGATG 1292
 QY 1300 GTGTGGGTTGATCTTTGAGATTTAACTGAGAGGTTGCTGCGCAAGAGATGAGGTG 1359

Db 1293 AACTGGGCTGAGCCTTAACAAATTCATCCTGATGATCATCATGATGATG----- 1344
 Qy 1360 GAGTTTTCGGTCTTGGGTCGAGATTGAGACTTGACCTTGGGGTCGGGTCGAGATT 1419
 Db 1345 ----TCAATATCTTGGTTCGATTTAGGTTGGACCTTGGCTCGGTTAAAGATT 1400
 Qy 1420 TGGCCCGGGAAGATCTTGGTTTACTACCTGATGTTTGGACGGCGATGATGTTACT 1479
 Db 1401 TGGCCCGGCAAAAGATGGGATTTGGCTGAGTTTCATCTTGGTGGCTGAGTTGCTGAA 1460
 Qy 1480 GAGTTTGAATGGGAGCGGTCGATGTAAGCGGTTGATCTTAATCTGAGAACTGAGGCTT 1539
 Db 1461 AGCTTCAATTTGCTCTTC---GAGAAATGATGATGATTTGCTGATGCTTAAAGATG 1517
 Qy 1540 TCTTCCGAGATGCTTATCTCTT 1562
 Db 1518 TCTCTCGAGATGAAAGATCTTT 1540

 RESULT 10
 US-10-163-198-34
 ; Sequence 34, Application US/10163198
 ; Publication No. US2003012645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebecca E. Cahoon
 ; APPLICANT: Elmer P. Heppard
 ; APPLICANT: No. US2003012645A1uhiro Nagasawa
 ; APPLICANT: Hajime Sakai
 ; TITLE OF INVENTION: Alteration of Embryo/Endosperm Size During Seed Development
 ; FILE REFERENCE: B81487 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,198
 ; PRIOR FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 60/295,921
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 60/334,317
 ; NUMBER OF SEQ ID NOS: 111
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 34
 ; LENGTH: 1930
 ; TYPE: DNA
 ; ORGANISM: Aquilegia vulgaris
 ; US-10-163-198-34

 Query Match 22.1%; Score 352; DB 15; Length 1930;
 Best Local Similarity 56.3%; Pred. No. 2.1e-108;
 Matches 786; Conservative 0; Mismatches 565; Indels 45; Gaps 5;

Db 573 GAAAGAAATAGAGAAAAATGATTAATGAGATTAATGCAATGAGGTTAAATGGGG 632
 Qy 614 CCGTTTTCGATGATGCTTAAACGGGTCGCTTAAACAACTGATGCTGAT 673
 Db 633 AAGTTAAGTTAAAGGATTAATACCTTGGCTTTAAACAAATGATGATGATGTT 692
 Qy 674 TCGAACAGATATAGCTTGAACAAAAACATGTTAGTACGT-----GAAATGG 724
 Db 693 TTGAAAGTTTATGATTTAAACAACTTAATGATGAGGTTTAACTTGAAGATGG 752
 Qy 725 TCGAAGAGTATATGATGCTCGAAGGTTGATGATGATGATGATGATGATGATGAT 784
 Db 753 TGAATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 812
 Qy 785 CGAGTTTATCTTCAAAAGCTCCGCTGATGATGATGATGATGATGATGATGATGAT 844
 Db 813 GCTGATGATTTGCAAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 872
 Qy 845 GATTGATCCGAGTTATATCCGAACCCGTTAATCAACCGGT----- 888
 Db 873 TTTTGTGAAAAATTTATGAGAAACAGAAACAGAGATTAATGATGATGATGATG 932
 Qy 889 --GATTTGCTGATGATTTGATGAGAGATTTGCTCTCCCTCATGATGATGATGAT 946
 Db 933 AAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 992
 Qy 947 CGAGCCGGAACATATCCGCTCTTGGAGATGATTTGAGAGAAACAGACAGATG 1006
 Db 993 GTATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1062
 Qy 1007 CGATCTTAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
 Db 1053 CAATCTCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1112
 Qy 1067 TACAAACAGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1126
 Db 1113 CCGATCTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1172
 Qy 1127 CTTCATCTCATATCTTAACGCTGATGATGATGATGATGATGATGATGATGATG 1186
 Db 1173 CCAACCTCTTATCTCAAGAGATGATGATGATGATGATGATGATGATGATGATG 1232
 Qy 1187 CATTCTATCATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1246
 Db 1233 CCTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1292
 Qy 1247 CGGAGGAGCAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1306
 Db 1293 CAGCTGGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1352
 Qy 1307 TTGATCTTGGAGTTTAACTGAGAGATGATGATGATGATGATGATGATGATGATG 1366
 Db 1353 CTAGGCAATGATTTAAACCGAGATGATGATGATGATGATGATGATGATGATG 1400
 Qy 1367 CGGTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1426
 Db 1401 GATTTATGAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1460
 Qy 1427 GGAAGAACTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1486
 Db 1461 GAAAGCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1520
 Qy 1487 AATGAGACCGTTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1546
 Db 1521 AATGAGATCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1577
 Qy 1547 AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1562
 Db 1578 AATGAGAGAGCTTT 1593

RESULT 11

US-09-349-385-9
 ; Sequence 9, Application US/09349385
 ; Patent No. US20020152495A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Ito, Toshio
 ; APPLICANT: Fromm, Michael
 ; APPLICANT: Meyerowitz, Elliot
 ; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
 ; FILE REFERENCE: MBI-0002
 ; CURRENT APPLICATION NUMBER: US/09/349,385
 ; CURRENT FILING DATE: 1999-07-09
 ; EARLIER APPLICATION NUMBER: 60/115,967
 ; EARLIER FILING DATE: 1999-01-15
 ; NUMBER OF SEQ. ID NOS: 13
 ; SOFTWARE: Patent In Ver. 2.0
 ; SEQ ID NO 9
 ; LENGTH: 1990
 ; TYPE: DNA
 ; ORGANISM: Pinus radiata
 ; FEATURE:
 ; OTHER INFORMATION: CDNA
 ; US-09-349-385-9

Query Match 21.3%; Score 348.2; DB 10; Length 1990;
 Best Local Similarity 55.5%; Pred. No. 4.3e-107;
 Matches 768; Conservative 0; Mismatches 588; Indels 27; Gaps 4;

QY 211 ATTCCGGGTCGCAAAAGGCTCCCTTGGTGAAGCATGTCATGTCAGACACTCTA 270
 DB 307 ATTCTGCGCCCAAGAGATGCGCATATTGGAAGTCTAATGACATGAGCGCTT 366
 QY 271 GCTCACCAGCAATCGCTGATGCAAGTGAAGAAATTCGAGCCAGAGGCTCATGCTTTC 330
 DB 367 CTCACGCGAAGCTGATCTCTGCTCGCTCATGCTGCGAAGAGAGCTCATGCTTTC 426
 QY 331 AGCTTAGAGAGAGCTCGGCTGATCGTCAAGTGAATCCGAGCTAGCGAAGAGATTCG 390
 DB 427 AGCTTAGGCTGACGCGCGCTGTTATGCTGAGATCCGAGTGGCGCCGAGTTCCTC 486
 QY 391 AATAGCCGGTCTTGGTGTATGACCGCGTTAAAGAAATCGGCTTACTCATGATGTTTAA 450
 DB 487 ACCGCCCTCACTTCGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 546
 QY 451 AGAGCAATGCTTTTGTGACACAGAGTGTACTGAGCAAGCGTTGCGGCTATCGCTTCG 510
 DB 547 AGGCTATGCTGCTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 606
 QY 511 AACCATCTCTTGTAGTCAAAACAAATCAGAAAGAGAGAGAGAGAGAGAGAGAGAGAT 570
 DB 607 GCTCATCTCTTGTGCTGCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 666
 QY 571 AGCCAGATGCTGATGTTTCTTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAT 627
 DB 667 GTGCAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
 QY 628 GAGTGTCTTAAACGAGGCTGCTTAAACATGATGATGATGATGATGATGATGATGAT 687
 DB 727 GCAACATGCTGAT 786
 QY 688 GAGCTGAAAAAACAATGTTG-----AGTTACGTAAATGTCGAAAGAGAT 738
 DB 787 GACATGCTCTCATGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 846
 QY 739 GATTGTCTCGAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 788
 DB 847 CAATCTGTTAGAT 906
 QY 799 CAAGAGCTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 858
 DB 907 CTTCGATTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 966
 QY 859 ATTATATCGAAGAT 906

DB 967 ATCATCGAAGAT 1026
 QY 907 GTCAAGCTTTTGTCTCTCCCTCCATGTTCAATTAATATCCAGACCGGAGATATGCGC 966
 DB 1027 GTTACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1086
 QY 967 GTTCTTTGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1026
 DB 1087 GTTCTCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1146
 QY 1027 CTGCTAGAGATGCTCTTCAATCCAGATATGATGATGATGATGATGATGATGATGAT 1086
 DB 1147 ATGAG 1206
 QY 1087 GTAGCGGAGAGATGAT 1146
 DB 1207 GTAGTGGGAT 1266
 QY 1147 GCTGTGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1206
 DB 1267 GAGTGTGAT 1326
 QY 1207 TTGGCGATTAAGAT 1263
 DB 1327 CTCTTACAGAT 1386
 QY 1264 ATGAT 1323
 DB 1387 ATGAT 1446
 QY 1324 AAACCTGAT 1383
 DB 1447 GCTCGAT 1506
 QY 1384 TTGAGACTTGAAGAT 1443
 DB 1507 CTGAT 1566
 QY 1444 ACTACGCTTATGTTTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1503
 DB 1567 GCCACTGTTATCTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1626
 QY 1504 GGTACGCGTGAAGAT 1563
 DB 1627 GAACTCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1686
 QY 1564 CCT 1566
 DB 1687 CAT 1689

RESULT 12
 US-10-163-198-22
 ; Sequence 22, Application US/10163198
 ; Publication No. US20030126645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebecca E. Cahoon
 ; APPLICANT: Elmer P. Heppard
 ; APPLICANT: No. US20030126645A1
 ; APPLICANT: Hajime Sakai
 ; TITLE OF INVENTION: Alteration of Embryo/Endosperm Size During Seed Development
 ; FILE REFERENCE: BBI487 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,198
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 60/295,921
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 60/334,317
 ; NUMBER OF SEQ. ID NOS: 111
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 22
 ; LENGTH: 1934
 ; TYPE: DNA

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rmpb

Page 12

ORGANISM: Glycine max
US-10-163-198-22

Query Match 19.9%; Score 316.6; DB 15; Length 1934;
Best Local Similarity 55.1%; Pred. No. 2.6e-96;
Matches 750; Conservative 0; Mismatches 569; Indels 42; Gaps 5;

254 TCATGTCAAGCACTCTAGCTCACCAGCAATCGCTGATGCACTGAGAAATTCGAGCCA 313
255 TCTTACTGCTCCACGCGCTCACCCTGCTTATCAATCGCCCTGATTAACACGCG 314
314 AGAGCTCATGCTTTTCACTTAGAGAGACTCGCGATGCTGACGTAATCCGAGC 373
315 AAAAAGTATGCTTTCTCCATCGGTTTAACTCCCTGCTATCTCCAGGAGACGAGA 374
374 TAGCGAAAGATTTCTGATATGACCGGTTTCTGCTGATCGACCGGTTAAAGATCGCTT 433
375 CCGTAAAGAGATTTCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 434
434 ACTCATGATGTTTAAAGAGAGATTTGTTGACACAGAGAGAGAGAGAGAGAGAGAG 433
435 ATGACCTTCTCTCCACCGCGCAATGGGTTTGAACGCTATGGGAGAGAGAGAGAGATT 494
494 TTGCGCGATGCTTCAACCATCTCTTATGACAAACAAATGAGAGAGAGAGAGAGAGC 553
495 TGAGAGAGATCTCAGCCCTTATCTCTCCGAGAGAGATCAGCGGCTTGAAATCCT 554
554 AACGAGAGATCTCAAGCAGATGTTGAGTTTCTTGAAGAAACAGAGATGTT---AAGC 610
555 TCAGAGAGAGAGTTGATTAATAATGTTGAACAAGTTAAACCAATGAGAGAGAGAGC 614
611 AACCGTTTGTGCTGAGTCTTAAACGCGCTTAAACCAATGAGAGAGAGAGAGAGAGC 670
615 AACATGTTGAGTTTAAAGAAATTTACACTTATGCTTGAACAAATGATGATGAGAGC 674
671 TATTCGAGACAGATGATGAGCTT---GAAAAAACCATGTTGATGATGATGATGATG 727
675 TCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 734
728 AAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 787
735 GTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 794
788 AGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 847
795 GGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 854
848 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 887
855 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 914
888 -TGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 946
915 ATGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 974
947 CCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1006
975 GTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1034
1007 CCGTCTTAAATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1066
1035 CATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1094
1067 TACAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1126
1095 CAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1154
1127 CTTCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1186
1155 GGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1214
1187 CACTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1243

1215 CGCTACTCTGCTGAGGCTGCGCTTGTGTCAGACGATACCGTGGGCGCAAGCAGTGA 1274
1244 TTCCGAG 1303
1275 TTCCCAAG 1334
1304 GGGTGTATGCTTTGAGATTTAACTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1363
1335 GGGCGAG 1382
1364 TTTCGATGCTTTGAG 1423
1383 TGAGCATATGAG 1442
1424 CCGGAG 1483
1443 CTGGAG 1502
1484 TTGAATGAG 1543
1503 TTCAATGAG 1562
1544 GCGAGATGAG 1584
1563 TGAGATGAG 1603

RESULT 13

US-10-163-198-40
Sequence 40, Application US/10163198
Publication No. US20030126645A1
GENERAL INFORMATION:
APPLICANT: Rebecca E. Cahoon
APPLICANT: Bimer P. Heppard
APPLICANT: No. US20030126645A1uhiro Nagasawa
APPLICANT: Hajime Sakai
TITLE OF INVENTION: Alteration of Embryo/Endosperm Size During Seed Development
FILE REFERENCE: B1487 US NA
CURRENT APPLICATION NUMBER: US/10/163,198
CURRENT FILING DATE: 2002-06-05
PRIOR APPLICATION NUMBER: 60/295,921
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 60/334,317
PRIOR FILING DATE: 2001-11-28
NUMBER OF SEQ ID NOS: 111
SOFTWARE: Microsoft Office 97
SEQ ID NO 40
LENGTH: 1687
TYPE: DNA
ORGANISM: Alstroemeria caryophylla
US-10-163-198-40

Query Match 18.2%; Score 290.6; DB 15; Length 1687;
Best Local Similarity 54.1%; Pred. No. 1.7e-87;
Matches 712; Conservative 0; Mismatches 579; Indels 24; Gaps 5;

257 TGTCAAGCACTTACTAGCTACCGAGAGATGCTGATGCAAGTGAAGAAATTCGAGCAAGA 316
193 TCTCCGCGCCCTCCGCGCCACCGCTCCGCGCGCGCTCCTCCCTCCGCGCGCTCC 252
317 GGTCTATGCTTTTCACTTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 376
253 CCTTCCTCTCTCTCTCCCTGCGCTCAACCGCTTCACTTCTCTCTCAACCTTCAACG 312
377 CCAAGAGATGCTGATGAG 436
313 CCAAGAGATGCTGATGAG 372
437 CACTGATGTTTAAAG 496
373 GCTTCCTCTCTCAACCGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 432
497 GCGGATGAG 556

Db	433	GCCGCATCTCCGCCACCCCACTCTTCAGCCCCAAGCCCTCTCCGCGCTCCGCCCCCTCC	492
QY	557	GACGAGTGATCTCAAGCCAGATGGTTGATTTCTTGAAAAAACAAGTATGTAAGAACCT	616
Db	493	GCCGCGCATCTCGGCGCTCCGCGCGCTCCCAACGTGCGCTCCCTCATAGGCCAACCGCG	552
QY	617	CTTTTGTTCGTAG--TTGGTTTAAACCGCGGTGCTTAAACAACATAGTGGCTCTGTAT	673
Db	553	AGGTCGAGATCAAGCGCCTCTCCACTTCGCTCTCCCTCAACAACGTATAGGCCAGCTGT	612
QY	674	TCGCAACAAGATATAGCTTGAAAAA--ACATGTTGAGTTACGTAAATGGTCGAAG	730
Db	613	TCGGCCGCGTCTACGACTTGGCCAACCGGGAAGCGCTCGAGCTCGAGGCGCTTGGTCAGCG	672
QY	731	AAGTTATGATTTGCTCGGAACGTTGAATTGGAATGATCAACTCTCTTGGGTATCGGAGT	790
Db	673	AGGGGTACGAGTGTGGGCGTCTTCAACTGGGGCGAACATTTCCGCTGTGCTGTGT	732
QY	791	TTGATCTCAAAAGACTCCGCTCTAGATGTTCCACACTGTACCAAGGTAAACCGATTG	850
Db	733	TTGACTTCCAGGGGGTCAAGCGAAGTGAAGGCCCTGTCAAGCGGATCAAGCTTTGG	792
QY	851	TATCCCGAATTATATCCGAACAC--GTATCAAAACGGTGAATTTGCTCTGTATTTCC	907
Db	793	TCGGCCGATATATGACAGACAGCACGCAAGCGGGCGAGGGCTCGTCAGCAACGAGC	852
QY	908	TCGAGTTTGTCTCCCTCCATGAGTTCAATAATTATCCGACCCGACATAATCGCCG	967
Db	853	CCGTAGACTTGTGACGCTCTGCTCCAGCGAAGACTCTCCGATTTCCACATGATGGCG	912
QY	968	TTCTTTGGAGATGATATTCAGAGAACAGACACATTTGCGGCTTTAATCGATGATCC	1027
Db	913	TTCTGTGGAGATGATCTTTCGCGGAGCGAATACGTTGCGCATCTGCTGAGTGAATCA	972
QY	1028	TCGCTAGATAGTCCCTTATCCAGATATGCAATCAAGGTTCAAAACGAGCGGATCAAG	1087
Db	973	TGGCGAGGATGGTGTGCAACCGGAAATCCAGCCAAAGCTCAACCGAATTCAGCGCG	1033
QY	1088	TAGTCGGAATTCAGAGACCCCTAGATGAATCTGACTTGAATTCATATCTTAACG	1147
Db	1033	TTGTGGCGGTGAGAGAGTCCGATGGCCGACGCCGACGTGCGCAACTCTCTTAACCTCAAG	1092
QY	1148	CTGTGGTGAAGAAGTATTAGAGCTTCACTCCACAGGCCACTTGTATCATGGGCGGAT	1207
Db	1093	CCATGTCAAAGAGTCGCTGAAGATGACCCCCCGGCGCGCTGCTCTCTGAGCTCGCC	1155
QY	1208	TGGCATTAAACAGACAGATCGTTGATGTCGTCTTGTTCCGCGAAGGACACACAGCATGG	1267
Db	1153	TCGCAGTCCATACGTGACGTCGCGGGGCACTTGTCCCGCCGACGACCGGATGG	1217
QY	1268	TGAACATGTGGGCCGTATCGCATGATTCACACAGTGGGTTATCTTTGAGTTTAAAC	1332
Db	1213	TGAACATGTGGGCCATAGCGCACGACGCGAATCTGGCCGAGCGGAGCGGATTCACC	1277
QY	1338	CTGACAGGTTGCTGGCAAAAGAGGAGGTGAGATTTTCGGTTCTTGGGTGCGATTGGA	1387
Db	1273	CGAAGAGTTTGTG-----GAGCAGGATGTAGCATTTCTGGGCTCGAATCTCC	1322
QY	1388	GACTTGCACCTTTCGGGTGGGTCGTGGATTGGCCCCCGGAGAAATCTTGGTTTACTA	1447
Db	1321	GGCTGGCGCGTTTCGGGTGCGGAGAGAGGATGTGTCCCGCAAGCGAATGGGAGCTGGGCA	1380
QY	1448	CGGTTATGTTTGAACGCGGATGATNTTACATGAGTTTAAATGGGGAACGTCGATGTGA	1507
Db	1381	CCGCGCATCTCTGAGCTGCGCTCACTGCTTCAAGACTTCAAGTGGGTGGCTTCGGA---CA	1437
QY	1508	ACGGCGTTGACTTATCTGAGAACTGAGGCTTTCTTTCGAGATGGCTAATCTCTT	1562
Db	1438	ATGGCGTTGATCTCTCGAAAACTTGAAAGATGTCCCTTGAGATGAAAGTCCCTCT	1492

Query Match	18.0%	Score 286	DB 15	Length 1545
Best Local Similarity	54.1%	Pred. No. 6	1e-86	
Matches 692	Conservative	0	Mismatches 550	Indels 36
				Gaps 4
US-10-163-198-96				
US-10-163-198-96				
Sequence 96, Application US/10163198				
Publication No. US20030128645A1				
GENERAL INFORMATION:				
APPLICANT: Rebecca E. Cahoon				
APPLICANT: Elmer P. Heppard				
APPLICANT: No. US20030126645A1				
APPLICANT: Hajime Sakai				
TITLE OF INVENTION: Alteration of Embryo/Endosperm Size during Seed Development				
FILE REFERENCE: B01487 US NA				
CURRENT FILING DATE: 2002-06-05				
PRIOR APPLICATION NUMBER: 60/295,921				
PRIOR FILING DATE: 2001-06-05				
PRIOR APPLICATION NUMBER: 60/334,317				
PRIOR FILING DATE: 2001-11-28				
NUMBER OF SEQ ID NOS: 111				
SOFTWARE: Microsoft Office 97				
SEQ ID NO: 96				
LENGTH: 1545				
TYPE: DNA				
ORGANISM: Zea mays				
US-10-163-198-96				
306 GCGAGCGCAAGAGCTCATGGCTTTGAGTGAAGAGAGCTCGCGTATCGTCACGTGCA	365			
255 CCGCGCGCGCGCGCTGTGCGCTTCTCGTGGCGCTCAAGCGCTCGTATGAGAGCA	314			
366 TCCGACGCTGAGCAAGAGATTTCTGAATAGCCCGCTTTTGTGATTCGACCGCTTAA	425			
315 GCGGACACGCGCGCGAGCTCTGCGCACGCCCGCTTCGCGACCGCGCGCTGAAGGA	374			
426 ATCGCGTTATCACTGATGTTTAAAGAGCAATGTTTGGACACACGCGTATTACTG	485			
375 CCGGCGCGCGCGCGCTCTTTCACCGCGCATGCGGCTTTGCGCGCGCGAGTACTG	434			
486 GCGAGCGCTTCGCGGATGCTTGAAACCATCTTTTATGACAAACAAATAGAAAGAC	545			
435 GCGCGCGCTTCGCGCATAGCTCGCGTACTTTCAGCCCGCGAGCGTGGCGCGGC	494			
546 CGAGAGCGCAAGAGATGATCTCAAGCCAGATGGTGAAGTTCTTGAAAAACAGATAG	605			
495 GCGCGCGCGCGCGCGCATCGCGAGCGCAAGCTCGCGACCTTCCGCGCGCGCGG	554			
606 TAACGACCGCTGTTGTTGCTGATGCTTAAACGCGCGTGGCTTAAACATGATGTCG	665			
555 ACGAGAGTCT---GTGATGCGCGCGTGTCCACGCGGCGATCCCTGACCAAGTCATGCG	611			
666 CTCTGTATTCGCAAGAGATGAGCTTAAA---AAACCATGTTGATTAAGTGAAT	722			
612 CACCGTTCGCGCGCGCTACACACCGCCAGCCCGAGGCGCGAGCTGAGAGAT	671			
723 GGTGCAAGAGATGATGATTTGTCGCAAGCTGATGATGATGATGATGATGATGATGAT	782			
672 GGTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	731			
783 ATCGAGATTGATCTCTCAAGAGCTCGCGTCTAGATGTTCCACACTGTATCAAAAGTAA	842			
732 CAGTGGCTGACCTCAGAGCGCTCAGAGCGCGTCAAGAGCTGTGGAGAGTCAA	791			
843 CCGGTTGATCCCGATTAATATCCGACACCGTAAATCAACCGGTGATTTGCTCGTGA	902			
792 CGGTGTCGCGCAAGTATCAAGAGCACAGGACAGAGAGAGAGAGAGAGAGAGAGAG	851			
903 T-----TTGTCGACGTTTGTCTCTCCCTCCATGCTTCAAGTAAAT	944			
852 GCGCGCGCGCGCGAGATTCGTGACGCTTGTGCGAGCTGAGAGCGAGAGAGAGCT	911			
945 ATCGAGCCGCAATATCGCGCTTTTGGAGAGATGATTTAGAGAAACAGACAGT	1004			

Db 912 GTGGATCTCGACATGATCGCTGCTCTCTGGAGATATCTTTCAGAGGACCGACAGGT 971
 Qy 1005 TGGGATCTTAATCGATGATCTTCGCTAGAGATGCTTTCATCCAGATATGCAATCAAC 1064
 Db 972 GGCGATCTGTGTGAGTGGGATGAGCGGAGATGCTGTCACCCGGGCAATCCAGTCCA 1031
 Qy 1065 GGTACAAAAGAGTGTGATCAAGTATCGGAAATCAAGCCCTAGATGAATCTGACT 1124
 Db 1032 GGCACAGCGAGAGCTGACCGCTGTGTGGCCGCGCCGCTGTTGCAAGCCAGCT 1091
 Qy 1125 GGTTCACCTTCATATCATCGAGCTGTGTGAAGAGATATTAAGCTTCATCTTCAGG 1184
 Db 1092 GGCCCGCTGCTGCTACCTGACGCGGCTGTGAAGAGACGCTCCGCTGACCCCGCGGG 1151
 Qy 1185 CCCACTTTCATATGAGCCGCTTGGCCATTAACAGACAGATCGTTGATGCTGTCTGT 1244
 Db 1152 CCCGCTGCTCTGTGGGCGCGCTGCGCTGCGAGCGGCTGTGCGCGGCACTGTGT 1211
 Qy 1245 TCCGACAGGACCAAGCATGTGAACATGTGGCGGTATCGCATATCCACAGCTGTG 1304
 Db 1212 CCCCGCGGACACAGGCAATGTGCAATGTGGCCATGCGCAGACCCCGCGGTGTG 1271
 Qy 1305 GATTATCTCTTGAAGTTAACTGAGAGGTTCTGTGCAAAAAGAGTGAAGTGTGAGT 1364
 Db 1272 GCCGAGCCCTCCGCTTTCGCGCCGAGCGGTGAG-----GTGAGAGAGCT 1319
 Qy 1365 TTCGTTCTTGGGTGAGATTGAGACTTGCACCTTTCGAGTGGGCTGTCCGATTTGCC 1424
 Db 1320 GAGCTGTGCGCGGCGGCACTCCGCTGCGCCCTTCGCGCCGCGCGCGGTGTGCC 1379
 Qy 1425 CGGAGAGATCTTGTGTTACTACCTGATGTTTGAAGCGGATGATTAATGAGTT 1484
 Db 1380 GGGCAAGACGTGGGCTGCGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1439
 Qy 1485 TGAATGGGAGCCGTCGATGATGAAGGGGTTGACTATCTGAGAACTGAGGCTTCTTG 1544
 Db 1440 CCGGTGGGCGCGCGCGAGCGCGCGGCTGCACTGTGCGAGGCGCTTCGCAATCTCCT 1499
 Qy 1545 CGAGATGAGCTAATCTCT 1562
 Db 1500 GGAGATGAGAGAGCCCT 1517

RESULT 15
 US-10-163-198-94
 ; Sequence 94, Application US/10163198
 ; Publication No. US20030126645A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rebecca E. Heppard
 ; APPLICANT: Elmer P. Heppard
 ; APPLICANT: No. US20030126645A1unh1ro Negasawa
 ; APPLICANT: Hajime Sakai
 ; TITLE OF INVENTION: Alteration Of Embryo/Endosperm Size During Seed Development
 ; FILE REFERENCE: B1487 US NA
 ; CURRENT APPLICATION NUMBER: US/10/163,198
 ; CURRENT FILING DATE: 2002-06-05
 ; PRIOR APPLICATION NUMBER: 60/295,921
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 60/354,317
 ; PRIOR FILING DATE: 2001-11-28
 ; NUMBER OF SEQ ID NOS: 111
 ; SOFTWARE: Microsoft Office 97
 ; SEQ ID NO 94
 ; LENGTH: 1758
 ; TYPE: DNA
 ; ORGANISM: Zea mays
 ; US-10-163-198-94
 Query Match 17.8%; Score 283; DB 15; Length 1758;
 Best Local Similarity 55.1%; Pred. No. 7e-85;
 Matches 709; Conservative 0; Mismatches 530; Indels 48; Gaps 6;

Qy 306 CGAGCCAGAGGCTCATAGGCTTTCAAGCTTAGAGAGACTCGCGATATCTGACGTGCA 365
 Db 444 CGGCGGCAAGGACATGATGCGCTTCTCGGTGGGCTGACCCGTCTGTGTGTGAGCA 503
 Qy 366 TCCGAGCTAGCGAAAGATTCGAAATAGCGGGTTTGTGTGATGACCGGTTAAGA 425
 Db 504 GCGCATAGCGAGCGGAGATCTGCGAGCCCGGCTTGGGCAACCGCCCGTCAAGA 563
 Qy 426 ATCGCTTACTCATGATGTTTAAAGAGCAATGTTTTCACCAAGG--TGTTTA 482
 Db 564 CGGCGGCGGCAACGCTGCTTTCACCGGCAATGAGGCTTCGCGCTTCGAGAGAGCGCA 623
 Qy 483 CTGGGAAAGCTTGGCGGTATGCTTGAACCATCTTTAGTAAAGCAATCAAG 542
 Db 624 CTGGGCGGCGCTCGCGCGCTGCGGCAACCATGTTGGCCCGCGCGCGCTGGCGG 683
 Qy 543 AGCCGAGACGCAACGAGATCTCAAGCCAGATGAGTTGATTTCTTGAAGAACAGAG 602
 Db 684 TCCCGGACACACCGCGCTCATGCGGAGGCAATGATGCGGCAATCGCGCTGCGAT 743
 Qy 603 TAGTAAAGAACCTGTTTGTGTGAGT---TGTTAAAGCGGCTGCTTAAACAT 659
 Db 744 GCGCGGCGCACGCGAGGCTCTCAAGCGGCTGTGATGTGATCTTCAACACAGT 803
 Qy 660 GATGCTCTGTATTGCGACAAGATATGAGCTTGAAGAAAAACATGTTGAGTTAGTGA 719
 Db 804 CATGCGACCTGTTTGGCAAGCTTGAAGAGCTGAGAGCGGAGGCGCCCTTCTGGA 863
 Qy 720 A--ATGTCGAAAGAGTTAGATTTGCTCGAAAGTTGAATTGACATGACCTTCC 776
 Db 864 CGAGATGAGGCGGAGGCTGACGACCTCTGCGGACAGTTCAACTGAGTCACTGCTGC 923
 Qy 777 TTGGCTATGAGGTTGATCTCAAGAGCTCGGCTGATGTTTCCACTGTACCAA 836
 Db 924 ATTGCTCAAGATCTGACCCCGGCGGTGCGCGCGGTGCAACAGGCTGTGCGAA 983
 Qy 837 GGTAAACGGTTTGTATCCCGATATATTCGAAACCGTAATCAACCG-----887
 Db 984 GTGCAATGCTGTGTGCAAGATCATCTTGAAGCAAGAGGCGCGCGCAATGAGAG 1043
 Qy 888 -----TGATTTGCTCGGATTTGCGGAGCTTTTGTCTCCCTCATGCTTC 935
 Db 1044 AGTGTGGGAGATGAGGATGAGTATGCTGTCAGCTCTTCTTGGGCTGAGAGAG 1103
 Qy 936 AGATTAATTATCCGACCCGCAATAATGCGCGTCTTGGGAGATATATTCAGAGAAC 995
 Db 1104 GGAAGAGCTGTCAGATGCGGACATGATCGCTGTTCTTGGAGATGCTTTCAGAGCGC 1163
 Qy 996 AGACACAGTTGCGGTCTTAATCGAGTGAATCCCTGAGAGGTCCTTCAACAGATAT 1055
 Db 1164 CGACACCGTGGCATCTGATGAGAGTGGTCAATGAGAGATGCGCTGACCCGAGCAT 1223
 Qy 1056 GCATCAACGATCAAAAAGAGCTGATCAAGTATGCGGAAATCAAGAGCTTGAATGA 1115
 Db 1224 CAGGCGAAGGCCAGGCGGAGGCTGAGCGGCTGTGAGGATCGGCGCGCTGCGG 1283
 Qy 1116 ATCTGACTTGGCTTCACTTCAATATCTAAGCGGTGTGTAAAGAGTATTAAGGCTTCA 1175
 Db 1284 GCGGAGGTGCGGACCTTACCTTCAATCAATGATGATGTAAGAGAGCGTGGCATGCA 1343
 Qy 1176 TCTCCAGGCGCACTTATCATGAGCGCGCTTGGCCATAACAGACAGATGTTATGG 1235
 Db 1344 CCGCGAGCGCGCTGCTGTGTGTGGCGCGCTGCGCTGCGGCTGCGGCGG 1403
 Qy 1236 TGTCTTGTTCGCGAGGACACAGCAATGTGCAAAATGTGGCGGCTATCGATATCC 1295
 Db 1404 CCACCTGTGCTCCCGCGGACACAGCAATGTGCAAAATGTGCAAAATGTGCAAAATCC 1463
 Qy 1296 ACAGGTGAGGTATGATCTTGAAGTTTAACTGAGAGGTTGTGCAAAAGAGTGA 1355
 Db 1464 GCGATCTGGGCGAGCGGAGAGTTCCGCGCGGAGCGGTTCCAG-----GA 1511
 Qy 1356 GGTGAGTTTTCGCTTCTTGGGTGAGATTTGACCTTCCGGTTCGGTTCGTCG 1415

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rnpb

Page 15

Db 1512 GGAGGAGCTGAGCGTCTCTGGGAGCGACCTCCCGCTGGCCCCCTTGGCGCGCGGGCGCG 1571
Qy 1416 GATTGCCCCGGGAGAGATTGTTTACTACCGTTATGTTTGAACGGCGATGATGTT 1475
Db 1572 CGCCTGCCCCGGCAAGATCTGGCCCTGGCACCCACCACTCTGGGTGCCCCAGCTTCT 1631
Qy 1476 ACATGAGTTGATGGGAGACCGTCCGATGTTAOGGGTTGACTTATCTGAGAACTGAG 1535
Db 1632 GCAACAAGTTGAGTGGG-----CCGCCGGCGGGGGGTGACCTGTGAGAGGCTGAG 1685
Qy 1536 GCTTTCCTTGGAGATGACTAATCCTCT 1562
Db 1686 CATGTGCTGAGATGGCCACGCGCT 1712

Search completed: January 17, 2004, 06:00:01
Job time : 544 secs

OM nucleic - nucleic search, using sw model

11083.169 Million cell updates/sec

Title: US-10-022-025A-2

Sequence: 1 atggctacgaaactcgaaag.....tcgcgcgtagggcgagltta 1593

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

word size : 30

Total number of hits satisfying chosen parameters: 4

Minimum DB seq length: 0

Maximum DB seq length: 20000000000

Post-processing: Listing first 45 summaries

Database : GenEmbl.*

1: gb_bar:*
2: gb_hg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sfr:*
12: gb_ay:*
13: gb_vt:*
14: gb_vt:*
15: en_ba:*
16: en_fm:*
17: en_hum:*
18: en_in:*
19: en_fm:*
20: en_on:*
21: en_ov:*
22: en_ov:*
23: en_pat:*
24: en_pi:*
25: en_pi:*
26: en_ro:*
27: en_sfr:*
28: en_un:*
29: en_vt:*
30: en_hg_hum:*
31: en_hg_hum:*
32: en_hg_other:*
33: en_hg_hms:*
34: en_hg_hms:*
35: en_hg_hms:*
36: en_hg_hms:*
37: en_hg_hms:*
38: en_hg_hms:*
39: en_hg_hms:*
40: en_hg_hms:*
41: en_hg_hms:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

Result No.	Query Score	Match Length	DB	ID	Description
---------------	----------------	-----------------	----	----	-------------

C	1	976	61.3	55870	8	AC005819	AC005819	Arabidops
C	2	976	61.3	84835	8	AC006418	AC006418	Arabidops
C	3	232	14.6	740	8	ATH552847	AJ555847	Arabidops
C	4	141	8.9	504	8	ATH553770	AJ553770	Arabidops

ALIGNMENTS

RESULT 1
AC005819/c
10000000
00000000
P1N 11-VAR-2002

DEFINITION Arabidopsis thaliana chromosome 2 clone T3A4 map C1C06C03, complete sequence.

VERSION AC005819.3 GI:20197448
KEYWORDS HTG.

ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta

PREFERENCE
1 (bases 1 to 55870)
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis

Rounsley, S.D., Lin, X., Kall, S., Snee, I.F., Fujii, C.I., Masou, I.M.,
 Shen, M., Ronning, C.M., Fraser, C.M., Somerville, C.R. and Venter, J.C.

REFERENCE 2 (bases 1 to 55870)
AUTHORS Lin, X.
TITLE Direct Submission

JOURNAL
Submitted (09-MAR-2000) The Institute for Ge
Medical Center Dr., Rockville, MD 20850, USA

AUTHORS	Town, C.D. and Kaul, S.
TITLE	Direct Submission
	0000

COMMENT
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
On Apr 18, 2002 this sequence version replaced gi:6598489.

```

FEATURES
  source
    1..55870
    /organism="Arabidopsis thaliana"

```

```

/monoculture="genomic"
/cultivar="Columbia"
/db xref="taxon:3702"

```

```

/chromosome="2"
/map="C1C06C03"
/clone="T3A4"

```

```
misc_feature
1. .14747
/note="overlap with BAC clone F13A10
(AC006418.70079 84825) "
```

gene 92. .1276
/gene="At2g46630"

mRNA

```
/gene="At2g46630"  
/codon_start=1
```

```

/protein_id="AAC69930.1"
/db_xref="GI:3831447"

```

[illegible]

HDYNQNHNYNQHSYNQNHQGNPKMCHQPSSSDSENIMSTRVITIAENKGAWM

EILSPQGNCTGGCTGSSRSVHGCTGCKGRRLQSSSSSSDGEQKKTKTKVPRKAN
 SNLEPMKAFPAHNSNVQMINSIYVNSTASHHDPGVHLKISRRKPSDGNFHSVDIGNNGY
 TN"
 repeat_region 723..810
 /rpt_family="CACCAT"n
 repeat_region 1345..1382
 /rpt_family="AT_rich"
 repeat_region complement(1904)..1925
 /rpt_family="AT_rich"
 gene 3740..5558
 /gene="At2g46640"
 /note="synonym: T3A4.2; predicted by genscan and
 genefinder"
 mRNA join(13740)..4072,4525..4839,5083..5215,5407..>5558)
 /gene="At2g46640"
 CDS join(3740)..4072,4525..4839,5083..5215,5407..5558)
 /gene="At2g46640"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAC69921.1"
 /db_xref="GI:3831438"
 /translation="MYKSIDGVARNKKRNEGISEIKTKATILDQVGLVDALDNWF
 DGLVLTGFPGFDLTAKFEEDMDDECGSGLDLYVVDGSIINQVQESPLISVANK
 FKHEDHDGADLTDFHSGISIKTAETPAVTAALAEALIEPQKKTTLALFEMEDHDGYTWT
 HCKKNDNLDGEEVRYKONGSKISKRFSPVKKLVMSKEENDLRIPIKRMQMI
 KMKLKKHHPVDATKALKDKDVPYKPRNCEALSHSLNLQICQVGIWMTVEVYCELL
 TNLTKLVLPVPMHVLVFYIYICFLIITYINASP"
 gene 6216..7156
 /gene="At2g46650"
 /note="synonym: T3A4.3; supported by full length cDNA:
 Ceres:3743"
 mRNA join(6216)..6447,6552..6618,6704..7156)
 /gene="At2g46650"
 CDS join(6370)..6447,6552..6618,6704..6957)
 /gene="At2g46650"
 /codon_start=1
 /product="putative cytochrome b5"
 /protein_id="AAC69922.1"
 /db_xref="GI:3831439"
 /translation="MANLISFDVAKHKCRKNDWILHGKVVYDSTFMDHPPGDNLV
 LAYTGDAIDEDVYHNSKDAKELMKKTCIGDVQSTVPVYQYIPWEEKSTAAETT
 KESGSKLIIYIPLIIIGVAPALFYNNK"
 repeat_region complement(7227..7247)
 repeat_region /rpt_family="AT_rich"
 repeat_region 7693..7715
 /rpt_family="AT_rich"
 repeat_region complement(7799..7822)
 repeat_region /rpt_family="AT_rich"
 gene complement(8165..9980)
 /gene="At2g46660"
 /note="synonym: T3A4.4"
 mRNA complement(join(8165..8782,9006..>9980))
 /gene="At2g46660"
 CDS complement(join(8165..8782,9006..9980))
 /gene="At2g46660"
 /codon_start=1
 /product="putative cytochrome P450"
 /protein_id="AAC69923.1"
 /db_xref="GI:3831440"
 /translation="MNTKESLILPALSSKSVLSQNTLAFSLIANTVITLISLFM
 TYGPGANKKYLFGRLISGTSYKTVITPQPKRPYVSGMSLTAARRIADAERK
 GARRLAFSLGDFRVLVTCNPVAKELNSVFAFRPKESATSLMKRAIGRPHGV
 YWYTKLRISNHLSTKQIKRAETQRRVLTSSQVLELKQSSNECFVREILTKYASNL
 NMGVSVFGQSEYELKSKHVELREMTVEBGVDLDTLMTLHPLMEFQRLKSTSL
 VAPYNAFVRIILSEHNQGTDLPRDFVDTLSTHSDKLDSPDIIVALEMTIPFGIDTHT
 VALLFVILARWILHPDMQSTVQNELDQVGSRALDESDLASLPYLAVALYEVLRITL
 PEPQLISMARLITDITVDGRIVPAETGTMVWMAVSHDPHYVQDLEFKERVAKEKE
 GREFPELSDLELAFSGSGRRICPGKNLGFTTWMFTNMLHEBEMPSDGNQVDLS
 EKRLISECANPPIPALRRRRS"
 repeat_region 10124..10180
 /rpt_family="AT_rich"
 repeat_region complement(14979..15002)
 /rpt_family="(A)n"

repeat_region
15546..15568
/rpt family="AT_rich"
repeat_region
17214..17366
/rpt family="AT_rich"
repeat_region
17478..17503
/rpt family="(A)n"
repeat_region
complement(1752..17543)
/rpt family="AT_rich"
repeat_region
complement(17647..17688)
/rpt family="AT_rich"
repeat_region
complement(17874..17894)
/rpt family="AT_rich"
repeat_region
complement(18001..18060)
/rpt family="AT_rich"
repeat_region
complement(18152..19796)
/gene="At2g46670".
/note="synonym: T3A4.5; predicted by genescan and
genefinder"
mRNA
complement(join<19152..19310,19344..>19796))
/gene="At2g46670"
CDS
complement(join(19152..19310,19344..19796))
/gene="At2g46670"
/codon_start=1
/product="hypothetical protein"
/protein_id="AAC69924.1"
/db_xref="GI:3831441"
translation="MYLTCRFEESKSAEKAVALLESTSGEPKTPTESHEKTRKVRSD
QGSSTSSNENIGSSSVSPNVLQSLVTWNKQDSPLPVSNREKAASKVEVGSQS
TNBTAAGSSSTKPKPEBSAKRMRSQREALAKPKFKDKDCPFKHLKOIQDOT
DSMEYKNGCYVRYSQRKLAERPRVYKQCFRTVNSDASTKS"
19989..21343
/note="molecular marker Atb7"
misc_feature
complement(20037..21512)
/gene="At2g46680"
/note="synonym: T3A4.6; identical to SP:P46897; supported
by cDNA: gi15027938.gb AY045826.1"
mRNA
complement(join(20037..20039,20111..20756,20977..>21512))
/gene="At2g46680"
CDS
complement(join(20340..20756,20977..21336))
/gene="At2g46680"
/codon_start=1
/product="homeodomain transcription factor (ATHB-7)"
/protein_id="AAC69925.1"
/db_xref="GI:3831442"
translation="MTEGEGSPAMASAEPIITMKKKNNKNNKQRFDEQIKSL
EMAPSPRPSPRKKVQLARELGLQPROVAIINQNKRAWKSKELETJNLRQYN
LASQPSIKKKKQALVSELRQLEKTEOKTKCEERQSGDQAVALLSTHSEEN
NRKRKEPRFEMEKDKRGHGVCMCHDYEDDDNGSNIRGRYFGFEEEPHLIN
IVEPADSLCTSSDWMRFKSDTTLTLDQSSNNTPMDPWS"
repeat_region
complement(2636..22664)
/rpt family="AT_rich"
repeat_region
23537..23571
/rpt family="AT_rich"
repeat_region
23760..23783
/rpt family="AT_rich"
repeat_region
24595..24656
/rpt family="(CAAA)n"
repeat_region
complement(25135..25155)
/rpt family="AT_rich"
repeat_region
25903..25923
/rpt family="AT_rich"
repeat_region
complement(26667..26771)
/rpt family="AT_rich"
repeat_region
complement(26800..26901)
/rpt family="R:ATREPI ATREPI ATREPI dispersed repeat from
GB1D86608 a consenus."
30015..30038
/rpt family="(TAAA)n"
repeat_region
complement(31911..31933)
/rpt family="AT_rich"
repeat_region
complement(32444..33466)
/rpt family="AT_rich"
3578..35832
gene

Query Match 61.3%; Score 976; DB 8; Length 55870;
 Best Local Similarity 100.0%; Pred.No. 0; Mismatches 0; Indels 0; Gaps 0;
 Matches 976; Conservative 0; Mismatch 0; Indel 0; Gap 0;

QY 1 ATGGCTACGAACTCGAAGAGCTCTTATCTTTCGCTTTTTCGCAATGACGCGTTCTA 60
 DB 9980 ATGGCTACGAACTCGAAGAGCTCTTATCTTTCGCTTTTTCGCAATGACGCGTTCTA 9921
 QY 61 AGCAAAACCACTTTCCTTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 120
 DB 9920 AGCAAAACCACTTTCCTTTCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCTCCCT 9861
 QY 121 CTCTTCTTATGACACTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTTCGCGCGTTA 180
 DB 9860 CTCTTCTTATGACACTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTTCGCGCGTTA 9801
 QY 181 ATATCCGCTTCAATACCAAAACCGGAAACGTTATTCGCGGTCCAAAGGCTTCCCTTGGTT 240
 DB 9800 ATATCCGCTTCAATACCAAAACCGGAAACGTTATTCGCGGTCCAAAGGCTTCCCTTGGTT 9741
 QY 241 GGAAGCATCTCACTCACTGTCAGCACTCTAGCTCAACGACGATCGCTGATGAGCTGAG 300
 DB 9740 GGAAGCATCTCACTCACTGTCAGCACTCTAGCTCAACGACGATCGCTGATGAGCTGAG 9681
 QY 301 AAATTCGAGCCCAAGAGGCTCATGCGCTTTCAGCTTTCAGAGAGACTGCGCTGATGCTCAG 360
 DB 9680 AAATTCGAGCCCAAGAGGCTCATGCGCTTTCAGCTTTCAGAGAGACTGCGCTGATGCTCAG 9621
 QY 361 TGCATTCGAGCCCAAGAGGCTCATGCGCTTTCAGCTTTCAGAGAGACTGCGCTGATGCTCAG 420
 DB 9620 TGCATTCGAGCCCAAGAGGCTCATGCGCTTTCAGCTTTCAGAGAGACTGCGCTGATGCTCAG 9561
 QY 421 AAAGATCGGCTTACTCACTGATGTTTAAACAGAGCAATGTTTTCACCAACGCGTGT 480
 DB 9560 AAAGATCGGCTTACTCACTGATGTTTAAACAGAGCAATGTTTTCACCAACGCGTGT 9501
 QY 481 TACTGCGGAAACGCTTCCCGTATCGCTTTCGAAACCACTCTTTCAGTACAAACCAATCAG 540
 DB 9500 TACTGCGGAAACGCTTCCCGTATCGCTTTCGAAACCACTCTTTCAGTACAAACCAATCAG 9441
 QY 541 AGAGCCGAGACGCAACGAGATGATCTCAAGCCAGATGTTGATGTTTTCGAAAAACAG 600
 DB 9440 AGAGCCGAGACGCAACGAGATGATCTCAAGCCAGATGTTGATGTTTTCGAAAAACAG 9381
 QY 601 AGTAGTAACGAACCTGTTTTTTCGTAAGTTCCTTAAACGCGCTTCCTTAAACATC 660
 DB 9380 AGTAGTAACGAACCTGTTTTTTCGTAAGTTCCTTAAACGCGCTTCCTTAAACATC 9321
 QY 661 ATGAGCTCTGTAATCGGCAAGATGATGAGCTTGAATAAACAATGTTGAGTTACGTGA 720
 DB 9320 ATGAGCTCTGTAATCGGCAAGATGATGAGCTTGAATAAACAATGTTGAGTTACGTGA 9261
 QY 721 ATGAGCAAGAGGTTATGATTTGCTGGAACGTTGAATGAGCTGATCACTTCCTTGG 780
 DB 9260 ATGAGCAAGAGGTTATGATTTGCTGGAACGTTGAATGAGCTGATCACTTCCTTGG 9201
 QY 781 CTATCGAGGTTGATTCCTCAAGACTCCGCTGTAAGTTCGACATCGTACCAAGGTA 840
 DB 9200 CTATCGAGGTTGATTCCTCAAGACTCCGCTGTAAGTTCGACATCGTACCAAGGTA 9141
 QY 841 AACCGGTTTGTATCCCGGATTATATCGAACAACCGTAATCAACCGGTTATTCCTCGT 900
 DB 9140 AACCGGTTTGTATCCCGGATTATATCGAACAACCGTAATCAACCGGTTATTCCTCGT 9081
 QY 901 GATTTTCGTCAGCTTTTCTCTCCCTCCATGCTTCAATAATATCCGACCCGAGATA 960
 DB 9080 GATTTTCGTCAGCTTTTCTCTCCCTCCATGCTTCAATAATATCCGACCCGAGATA 9021
 QY 961 ATCCGCTTCTTTGGG 976
 DB 9020 ATCCGCTTCTTTGGG 9005

RESULT 2
 AC006418/c
 LOCUS
 DEFINITION
 complete sequence.
 AC006418
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Arabidopsis thaliana (thale cress)
 HTG.
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosid 1; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 84825)
 Lin, X., Kaul, S., Sheng, T.P., Fujii, C.Y., Shen, M., Vanhaken, S.E.,
 Barnstead, M.E., Mason, T.M., Bowman, C.L., Bowman, C.M.,
 Benito, M.-I., Carrera, A.J., Creasy, T.H., Buell, C.R., Town, C.D.,
 Niernman, W.C., Fraser, C.M. and Venter, J.C.
 Unpublished
 2 (bases 1 to 84825)
 Lin, X.
 Direct Submission
 Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA
 3 (bases 1 to 84825)
 Town, C.D. and Kaul, S.
 Direct Submission
 Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
 Medical Center Dr., Rockville, MD 20850, USA, cdtown@igf.org
 On Apr 18, 2002 this sequence version replaced gi:6598561.
 COMMENT
 FEATURES
 source
 1..84825
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="2"
 /map="C106C03"
 /clone="F13A10"
 1..36153
 /note="overlap with BAC clone F11C10
 (AC006526:59064..95214)."
 1..36153
 /note="overlap with BAC clone F11C10
 (AC006526:59064..95416)."
 complement(1047..3029)
 /gene="At2g46480"
 /note="synonym: F13A10.1; predicted by genscan"
 complement(join(<1047..1332,1409..1518,1595..1997,
 2074..2431,2503..2828,2916..3029))
 /gene="At2g46480"
 complement(join(1047..1332,1409..1518,1595..1997,
 2074..2431,2503..2828,2916..3029))
 /gene="At2g46480"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAD20159.1"
 /db_xref="GI:4415929"
 /translation="MTDACCLKGNDKRVPRFGHTWTKAFNDTPENLHRSRLROEK
 RLEANEIMNDLSLOKLTAMARSRSDAPLGNITWKNERYRGSFEDMLRLMD
 OIIMARVSGAKPNNALHOEITQMLKAMBESETDIOEQLVDSIRMGQILA
 RAHOLYBECKLVNTLRAMLOVEDELENEOTYITPLQLASKALPDALICITWELN
 EHLIPRLPRNPRENLENPCLYHVALPSDNVLAASVNSTVNAADPSHVEFLA
 TKLNFAGASWFLNPPGEATIHQREDFTWINSYSPLISQESAMKKFYKTA
 RSEVSQSENKIRYPRYSMLNLRITIPRIPEKLEKILFVDDDVVQDLTPIMS
 IDLKGVMENPDPKCGWAVGNMIPDLKWKKNNTTEYHFQNLNERTIKLGTLP
 PGLIFVNLTOPLQKRMHLGLGIDYKGIIVKIKERSAVIHVGNHKKPWTWGISKTOP
 YWKVTNPDHPYIFTCRLPE"
 4060..4696
 /gene="At2g46490"
 /note="synonym: F13A10.2"
 4060..4696
 rRNA

```
CDS
/gene="At2g46490"
/100.4504
/gene="At2g46490"
/codon_start=1
/product="unknown protein"
/protein_id="A020160.1"
/db_xref="GI:4415930"
/translation="MAADGIFRSIFEGCISGIDSLERRPYHKNOCALHDKSGACK
NKNORRPPSCRRGSSSISPIRIRMSKSGNIMAMTSSSSSSSSSSSSSSSSSSSS
LASDLPVNNATIEERSSKQRLMTIDGSD"
repeat_region
complement(4483..4517)
/rpt_family="(CA)n"
complement(6756..6792)
/rpt_family="AT-rich"
complement(11207..13602)
/gene="At2g46500"
/note="synonym: F13A10.3; similarity to ubiquitin family
of proteins; supported by cDNA: gi_16930424_gp_A0419366
1_A0419366"
complement(join(11384..11764,11858..13377))
/gene="At2g46500"
complement(join(11384..11764,11858..13377))
/gene="At2g46500"
/codon_start=1
/product="expressed protein"
/protein_id="A020161.1"
/db_xref="GI:4415931"
/translation="MSAGVALSPVSEBPLMPLVRANSCDSYDDTIMYTLTPGS
VIMRVLBSIESVLRIOYRGVVRNOKLVFGRIASNSNMRDYVSEGNILH
LVKSLDLVLYKTCGKCRHFVERGNIGYVKQISKRGPVDEDEIIEGE
KLEDDSLINDICNDSDVHLVRSKAKVYKVPKNEPLSVAPQADKKGRAGSI
VPEKSLSPVYVNSKAKVPLVYKMIQASDGLKSGNSPVRSRGGTGAFFMGSG
KMGVGRKPIDEPMAENNPGLPLSPNGEGIKKGTXYRGALREVAITLHDKSR
KSMGRBISGAVPTPMTECHLPGPNPKIKITGSLQGFTERDSCCEMGLSP
VEVHKISVLDIRLAVADPNSNITLMTQDSGLVLPIDGTCIPSEFECEETEMLY
WPGARPYSAETQETIRSLDAEDIDLEKFMKMPATPQDTLITLMLKKYGERL
TAEIGTICRETLTKSKSLVEMVEADQAVLPQSEAFLEALSDVDVHLDHSDH"
gene
complement(14394..15192)
/gene="At2g46505"
/note="synonym: F13A10.20; supported by full length cDNA:
Cerec: 11006"
complement(<14394..>15192)
/gene="At2g46505"
complement(14456..15111)
/gene="At2g46505"
/codon_start=1
/product="Expressed protein"
/protein_id="AA015243.1"
/db_xref="GI:20197780"
/translation="MSLRKTIIDLRQTORATLSGLPSSSHISSASATANVPLG
RDLSTIPFQADQKLPDSTNLVDRSISSTIGQSLNTAAFSQSSRGYTNSSFLR
KLPVPHIHGMEELADYHQEMTRNLVMSLGFQIIVLKDIILFL"
gene
complement(15830..17530)
/gene="At2g46510"
/note="synonym: F13A10.4"
complement(<15830..>17530)
/gene="At2g46510"
complement(115830..17530)
/gene="At2g46510"
/codon_start=1
/product="putative bHLH transcription factor"
/protein_id="A020162.2"
/db_xref="GI:20197775"
/translation="NMMSDLGMDDEKSVSAVLSHLASDFLRANSNONLPLVMTG
DITLTKSLSLVDPNSENFSNMYAIFMQQTMSSRGQOVLAGGSCCEPVEEESVY
VRSYNNMGAAEETWDMRKRVLOKHLFGSGDEDVYALSLKVTATTEFLASVY
FRNHBEGGPGRCYSSGKXVLSDNSSDYCPSPAKAGCITTYMTDGVILE
LGSVMSLPENIGLVSVQALFMRVTQPVWVTSNNMGGLHKFGDLSLAHYPPK
LEVRNLDERTFQSWEGYNNKGPFGVTPQDRDVKYLVENVNVVDDNNKXTIEA
GSSVAASNRSTNTQOKESCTEKRPVSLAGAGIVVDEKPKRKGAPANGRE
PLNHVABERORREKLNQRFYALSVVPSKMDKSLGDAISYIKELQEVKXMBE
RVGTDKLSNNTITVEESPVDIQMNEEVVAVISPLDHPASRIIQAMRNNSVL
```

```
MEAKSLADNMFTFVYKSNNGSDPLTEKTLAAFPETRSIQPLPSSSSQVSGDI
"
repeat_region
/rpt_family="(GA)n"
complement(18392..18454)
/rpt_family="AT-rich"
21510..24428
/gene="At2g46520"
/note="synonym: F13A10.5"
<21510..>24428
/gene="At2g46520"
21510..24428
/gene="At2g46520"
/codon_start=1
/product="putative cellular apoptosis susceptibility
protein"
/protein_id="A020163.1"
/db_xref="GI:4415933"
/translation="WENNRFTVLSGCFNTLSPIEPPTTERALSDADQNTGL
AVRLVAPPAIDEQTRHAAVNFKNHRSRMPAGDSIPVDSKEQTKLIVSLM
LSASPIQSOSEALTVYKIDPPKAMPALPELIANTQNALAGDVSYNGILGAS
SIFKFSYEYTDALFVDLKYCLDNFAPALTEIFLKTSYSDASAGSGSPILKEPL
ESORLCCTIFYSINFDLPFEFEDHMKEMNGEPFKYSNYPALSTEBELVDDLR
AICENIHVYIEKXDEEBOGFLNFPASVMTLVDRKSPDQATTAIKPLTVST
SVHNLFRGNDVYIEICOSYIVPVVSLRVDETFEMNYIEFRDMESGDDVRRRI
AGEHLKGLATYKTOVTVSLEIQRLLSFSANPSANWDCOKATILVYSSTKAG
GASVETDIDVQNFANILPELQSRVNSFPMLKASLPLTFMRSHLKPSTKQAF
PELVFRLASNNVHVASACIEKLVLVKEGAGNVAAGDISPKPLQMTNLPAL
KPESEENQYLMKIMRVLGVADISAVAGCIGLTSISEVCNKPKNIFMNYLFE
SVAVOIRACRDISLSAFETSPSLQWILANDIEFPIYGFQILAQVLENRPL
SNVVOIFILLSPESMRSRGNVPALVRLQAFQOKPHVTCENRSLVYGLFEKVL
ASPSGQGFYILNTIENLDYSYALPYMGPMLLPTRONKVTVFQSLVPSL
FLYGRDQGVYVETMTPQNPITTAIVHFWIPVNLKIGSMETKLTAVATRLICTEP
ALHDSRAKMLKGLMDSIVTVVSPRDEPRVDEPMPEISENGYTPAFLKNAGCK
EDPFLKDIKDPQPLVASVSRSSASPERIPQIIIGENLBANOTALLQCNAYNCGIA
"
repeat_region
25088..25183
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.thaliana
microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
74A"
repeat_region
25090..25183
/rpt_family="(GA)n"
25090..25166
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.thaliana
microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
74A"
repeat_region
25090..25162
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.thaliana
microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
74A"
repeat_region
25090..25158
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.thaliana
microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
74A"
repeat_region
25090..25154
/rpt_family="Rf:gi|1100803|emb|X93607.1|AT74A.A.thaliana
microsatellite repeat region (clone 74A) #Rf:ATR0046|X93607
74A"
Query Match 61.3%; Score 976; DB 8; Length 84825;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 ATGCTACGAAGCTCGAAGCTCTTATCTTGTCCCTTTGTCCAAATGACAGCTCTA 60
DB 80058 ATGGCTAGCAAACTGAAAGCTCTTATCTTGTCCCTTTGTCCAAATGACAGCTCTA 79999
QY 61 AGCGAAACCAACTGCTTCTTCTCTCTGCGCCCTCAACATCACTGCTGCCATATCT 120
DB 79998 AGCGAAACCAACTGCTTCTTCTCTCTGCGCCCTCAACATCACTGCTGCCATATCT 79939
QY 121 CTCTCTTATGACCTATCCGCGTGAAGCTCTTGGGGAATACCTCTTGCGCGGTTA 180
DB 79938 CTCTCTTATGACCTATCCGCGTGAAGCTCTTGGGGAATACCTCTTGCGCGGTTA 79879
```

```

QY 181 ATATCCGGTTTCATACAAAACCGGAAAGCTTATCCCGGTCGAAAAGCTTCCCTTTGGTT 240
Db 79878 ATATCCGGTTTCATACAAAACCGGAAAGCTTATCCCGGTCGAAAAGCTTCCCTTTGGTT 79819
QY 241 GGAAGATGTCATCTATGTCAGACCTCTAGCTCACCGAGCAATGCTGATGCAAGTGAAG 300
Db 79818 GGAAGATGTCATCTATGTCAGACCTCTAGCTCACCGAGCAATGCTGATGCAAGTGAAG 79759
QY 301 AAATTCGGAGCCAAAGAGCTCATGCTTTCAGCTTTCAGAGAGACTGCGCTGATCGTACG 360
Db 79758 AAATTCGGAGCCAAAGAGCTCATGCTTTCAGCTTTCAGAGAGACTGCGCTGATCGTACG 79699
QY 361 TGCATCCGACGTAAGGAAAGATTCGAATAGCCCGGTTTTCGATGAGACCGGTT 420
Db 79698 TGCATCCGACGTAAGGAAAGATTCGAATAGCCCGGTTTTCGATGAGACCGGTT 79639
QY 421 AAAGATCGGCTTACTCATGATGTTTAAACAGACCAATTGTTTTCACGACACGCTGTT 480
Db 79638 AAAGATCGGCTTACTCATGATGTTTAAACAGACCAATTGTTTTCACGACACGCTGTT 79579
QY 481 TACTGGGGAACGCTTCCCGCTTATCGCTGCAACATCTCTTGTAGTACAAACAAATCAGA 540
Db 79578 TACTGGGGAACGCTTCCCGCTTATCGCTGCAACATCTCTTGTAGTACAAACAAATCAGA 79519
QY 541 AGAGCCGAGACGCAACGACGATGATCTCAAGCCAGATGATGTTTTCGAAAAACAG 600
Db 79518 AGAGCCGAGACGCAACGACGATGATCTCAAGCCAGATGATGTTTTCGAAAAACAG 79459
QY 601 AGTAGTACGAAACCTCTGTTTGTGCTGAGTGTCTTAAACGCGCTGCTTAAACAATG 660
Db 79458 AGTAGTACGAAACCTCTGTTTGTGCTGAGTGTCTTAAACGCGCTGCTTAAACAATG 79399
QY 661 ATGAGCTCTGATTCGGAACAAGATGATGAGCTTGAACCAATGTTGAGTACGAGAA 720
Db 79398 ATGAGCTCTGATTCGGAACAAGATGATGAGCTTGAACCAATGTTGAGTACGAGAA 79339
QY 721 ATGCTCGAAGAGGTTATGATTTGCTCGGAACGTTGATGATGATGATGATGATGATGAT 780
Db 79338 ATGCTCGAAGAGGTTATGATTTGCTCGGAACGTTGATGATGATGATGATGATGATGAT 79279
QY 781 CTATCGAGTGTGATCTCTGAAAGACTCGGCTGATGATGATGATGATGATGATGATGAT 840
Db 79278 CTATCGAGTGTGATCTCTGAAAGACTCGGCTGATGATGATGATGATGATGATGATGAT 79219
QY 841 AACCGGTTTGTATCCCGGATTAATTCGAAACACCGTATCAACCGGTTGATGCTCTGCT 900
Db 79218 AACCGGTTTGTATCCCGGATTAATTCGAAACACCGTATCAACCGGTTGATGCTCTGCT 79159
QY 901 GATTTCTGTCAGCTTTTCTCTCCCTCCATGATGATGATGATGATGATGATGATGATGAT 960
Db 79158 GATTTCTGTCAGCTTTTCTCTCCCTCCATGATGATGATGATGATGATGATGATGATGAT 79099
QY 961 ATCCGCTTCTTTGGG 976
Db 79098 ATCCGCTTCTTTGGG 79083

```

```

RESULT 3
LOCUS ATH552847/c 740 bp DNA linear PLN 29-MAR-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
345811.
ACCESSION AJ552847
VERSION AJ552847.1 GI:29368998
KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
1 Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,

```

```

TITLE
JOURNAL MEDLINE
PUBMED 12446535
REFERENCE 2 (bases 1 to 740)
AUTHORS Balzerque, S.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
http://dbgap.versailles.inra.fr/publiclines/. This sequence has
been generated in the framework of the French plant genomics
program "genoplante" (http://www.genoplante.com and
http://genoplante-info.infobiogen.fr).
FEATURES
source
1..740
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassilewskija"
/db_xref="taxon:3702"
/clone="345811"
/clone_1ib="Arabidopsis thaliana T-DNA insertion lines"
misc_feature
1..740
/note="T-DNA flanking sequence
left border"
BASE COUNT 192 a 172 c 197 g 179 t
ORIGIN
Query Match 14.6% Score 232; DB 8; Length 740;
Best Local Similarity 99.4% Pred. No. 5e-121;
Matches 332; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 200 CCGAAGACGTTATCCCGGTCGAAAAGCTTCCCTTTGGTGAAGATGTCATCATGT 259
Db 429 CCGAAGACGTTATCCCGGTCGAAAAGCTTCCCTTTGGTGAAGATGTCATCATGT 370
QY 260 CAAGCACTTACTCAACGAGCAATCGCTGAGCGAGTGAAGAAATTCGAGGCAAGAGGC 319
Db 369 CAAGCACTTACTCAACGAGCAATCGCTGAGCGAGTGAAGAAATTCGAGGCAAGAGGC 310
QY 320 TCATGCTTTCAGCTTGAAGAGACTCGGCTGATGATGATGATGATGATGATGATGATGAT 379
Db 309 TCATGCTTTCAGCTTGAAGAGACTCGGCTGATGATGATGATGATGATGATGATGATGAT 250
QY 380 AAGAGATTGTGAATAGCCCGGTTTTCGTCGATGACCGGTTTAAAGATCGCTTACTAC 439
Db 249 AAGAGATTGTGAATAGCCCGGTTTTCGTCGATGACCGGTTTAAAGATCGCTTACTAC 190
QY 440 TGAATGTTAAGAGAGCAATGTTTTCGACACAGAGTGTACTGCGAAGCTTGGCC 499
Db 189 TGAATGTTAAGAGAGCAATGTTTTCGACACAGAGTGTACTGCGAAGCTTGGCC 130
QY 500 GATTCGCTTGAACCATCTCTTGTAGTACAAACA 533
Db 129 GATTCGCTTGAACCATCTCTTGTAGTACAAACA 96

```

```

RESULT 4
LOCUS ATH53770/c 504 bp DNA linear PLN 29-MAR-2003
DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone
371810.
ACCESSION AJ553770
VERSION AJ553770.1 GI:29370237

```


Tue Jan 20 17:43:19 2004

us-10-022-025a-2.011go.rge

Page 6

KEYWORDS left border; T-DNA flanking sequence.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE
AUTHORS Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
Chauvin, S., Bechtold, N., Cruaud, C., Derose, R., Pelletier, G.,
Lepoint, J., Caboche, M., and Lecharny, A.
TITLE T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites
JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE 22363535
PUBMED 12446565
REFERENCE 2 (bases 1 to 504)
AUTHORS Balzergue, S.
TITLE Direct Submission
JOURNAL Submitted (21-NOV-2002) Balzergue S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
<http://dbsgap.versailles.inra.fr/publicines/>. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (<http://www.genoplante.com> and
<http://genoplante.info.inbio.gen.fr>).
Location/Qualifiers
1..504
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassilewskija"
/db_xref="taxon:3702"
/clone="372B10"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
1..504
misc_feature
/note="T-DNA flanking sequence
left border"
BASE COUNT 180 a 128 c 71 g 125 t
ORIGIN
Query Match 8.9%; Score 141; DB 8; Length 504;
Best Local Similarity 99.5%; Pred.No.1.le-68;
Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
1331 AGAGGTTCTGGGCAAAAGAGTGAGTGGATTTCGTTCTTGGGTGAGTTGAGAC 1390
|||||
497 AAGAGTTCTGGGCAAAAGAGTGAGTGGATTTCGTTCTTGGGTGAGTTGAGAC 438
|||||
1391 TTGACCTTTCGGGTCGGGTGCGATTGGCCCCGGGAAGATCTTGTTTACTACCG 1450
|||||
437 TTGACCTTTCGGGTCGGGTGCGATTGGCCCCGGGAAGATTTGGTTTACTACCG 378
|||||
1451 TTATGTTTGGACGGGATGATGTTTACATGAGTTTGAATGGGACCGTCCGATGGTAAAG 1510
|||||
377 TTATGTTTGGACGGGATGATGTTTACATGAGTTTGAATGGGACCGTCCGATGGTAAAG 318
|||||
1511 GCGTGACTTAT 1522
|||||
317 GCGTGACTTAT 306
|||||

Search completed: January 17, 2004, 08:54:47
Job time : 5880 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 00:04:29 ; Search time 3521 Seconds
(without alignments)
10996.033 Million cell updates/sec

Title: US-10-022-025a-2
Perfect score: 1593

Sequence: 1 atggctacgaactcgaaag.....tgcgcgtagcgagcttaa 1593

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 22781392 segs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

BST:
1: em_estba:*
2: em_estbm:*
3: em_estlm:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estlo:*
8: em_hrc:*
9: gb_estl:*
10: gb_est2:*
11: gb_hrc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfm:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vit:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rtd:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*
29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	566.4	35.6	712	29	BZ428523 BONHW74TF
2	527	33.1	784	28	BH492184 BOHFJ05TR
3	495.8	31.1	809	29	BZ428529 BONHW74TR
4	476	29.9	674	28	BH451907 BOGSO44TR

C	5	428	26.9	745	28	BH510149
C	6	421.8	26.5	727	28	BH527239
C	7	399.2	25.1	659	29	BZ466758
	8	391.8	24.6	776	28	BZ443616
	9	330.8	20.8	705	28	BH977803
	10	330.8	20.8	732	28	BH534160
	11	329.4	20.7	826	14	CB344213
	12	322.6	20.3	638	14	CB343993
	13	314.8	19.8	720	10	BG648529
	14	299.6	18.8	576	12	BH451899
	15	298.6	18.7	791	12	BG839564
	16	297.2	18.7	804	28	BH431683
	17	286.2	18.0	612	9	AV826713
	18	285	17.9	765	28	BH581933
	19	283.4	17.8	764	28	BH652028
	20	271.2	17.0	455	9	A1995549
	21	268.4	16.8	827	14	CB655569
	22	265.4	16.7	764	13	BO403090
	23	262.2	16.5	846	13	BG518647
	24	251.2	15.8	540	28	BH250286
	25	248.8	15.6	549	12	B1973185
	26	247.2	15.5	739	28	BH547739
	27	242	15.2	616	14	CA030582
	28	238.8	15.0	691	9	AW132420
	29	238	14.9	740	14	CB344274
	30	236.6	14.9	540	29	BZ436712
	31	233.2	14.6	489	10	BK345914
	32	231.2	14.5	641	12	EM063648
	33	229.6	14.4	635	14	CB628186
	34	229.2	14.4	656	9	AW622577
	35	224.4	14.1	835	10	BG447296
	36	220.4	13.8	454	10	BZ210324
	37	220.4	13.8	605	12	BZ226188
	38	217.6	13.7	810	10	BG583985
	39	213.6	13.4	712	14	CA783169
	40	212.6	13.3	468	9	AW586022
	41	212	13.3	653	10	BF595944
	42	211.8	13.3	495	10	BG580466
	43	211.8	13.3	537	9	AL656577
	44	207.6	13.0	531	10	BR191362
	45	207.4	13.0	524	10	BG998607

ALIGNMENTS

RESULT 1
BZ428523
LOCUS
DEFINITION
BONHW74TF BO.1.6.2 XB-loc Brassica oleracea genomic survey sequence.
ACCESSION
BZ428523
VERSION
BZ428523.1 GI:26671493
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 712)
AUTHORS
Town, C.D., Van Aken, S., Uteirbeck, T., Koo, H. and Frazer, C.M.
TITLE
Whole genome shotgun sequencing of Brassica oleracea
JOURNAL
Unpublished
Other GSSs: BONHW74TR
COMMENT
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.

Tue Jan 20 17:43:23 2004

us-10-022-025a-2.rst

Page 2

FEATURES
source

Location/Qualifiers
1. 712
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHM74"
/note="Vector: PHOS1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 168 a 179 c 177 g 188 t
ORIGIN

Query Match 35.6%; Score 566.4; DB 29; Length 712;
Best Local Similarity 87.2%; Pred. No. 5e-167;
Matches 621; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

251 CACTCATGTCAGCACTAGCTACACGAGCAATGCTGATGAGTGAATTCGAG 310
1 CTCTCATGTCAGCACTAGCTACACGAGCAATGCTGATGAGTGAATTCGAG 60
311 CCAAGAGGCTCATGCTTCAGCTTGAAGAGACTGCGTATGCTACATGCAATCCG 370
61 CCAAGAGGCTCATGCTTCAGCTTGAAGAGACTGCGTATGCTACATGCAATCCG 120
371 AGCTAGCGAAGAGATTCGAATGAGCGGCTTTTGTGATGAGCGGTTAAGAAATCG 430
121 AGCTAGCGAAGATTCGAATGAGCGGCTTTTGTGATGAGCGGTTAAGAAATCG 180
431 CTCTACTGATGATGTTAAACAGAGCAATGCTTTGACACACGCGTGTTAATCGG 490
181 CTCTACTGATGATGTTAAACAGAGCAATGCTTTGACACACGCGTGTTAATCGG 240
491 CGCTTGGCGATGCGTTCGAACCATCTTTTGAATGCAAAAGAAAGAGCGGAGA 550
241 CTCTGCGGATGCGTTCGAACCATCTTTTGAATGCAAAAGAAAGAGCGGAGA 300
551 CGCAAGAGAGATGATCTCAAGCAGATGATGATGATGATGATGATGATGATGAT 610
301 CACAGAGCGAGATGATCTCAAGCAGATGATGATGATGATGATGATGATGATGAT 360
611 AACCTGTTTGTGCTGAGTGTCTTAAACGCGGCTTTAACAACAGATGCTCTG 670
361 GCGTGTGTTTGTGCTGAGTGTCTTAAACGCGGCTTTAACAACAGATGCTCTG 420
671 TATTCGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 730
421 TTTTCGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 480
731 AAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 790
481 AAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 540
791 TTGATCTCAAGAGCTCCGCTGATGATGATGATGATGATGATGATGATGATGAT 850
541 TTGATCTCAAGAGCTCCGCTGATGATGATGATGATGATGATGATGATGATGAT 600
851 TATTCGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 910
601 TATTCGAGCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 660
911 ACCTTTTGTGCTGAGTGTCTTAAACGCGGCTTTAACAACAGATGCTCTG 962
661 ACCTTTTGTGCTGAGTGTCTTAAACGCGGCTTTAACAACAGATGCTCTG 712

RESULT 2
BH492184 784 bp DNA linear GSS 13-DEC-2001
LOCUS BH492184 BOHF0505R Brassica oleracea genomic clone BOHF05, genomic
DEFINITION survey sequence.
ACCESSION BH492184
VERSION BH492184.1 GI:17700288

KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

GSS.
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
eustroids; II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 784)
Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BOHF0505
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
source

Location/Qualifiers
1. 784
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHF05"
/note="Vector: PHOS1; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 204 a 189 c 211 g 180 t
ORIGIN

Query Match 33.1%; Score 527; DB 28; Length 784;
Best Local Similarity 84.2%; Pred. No. 1.5e-154;
Matches 622; Conservative 0; Mismatches 105; Indels 12; Gaps 2;

1 ATGGCTACGAAGCTCGAAGCTCTTAATCTTGGCTTTGTCGAATGAGCGTTCTA 60
736 ATGGCTACGAAGCTCGAAGCTCTTAATCTTGGCTTTGTCGAATGAGCGTTCTA 677
61 AGCCAAACCAACCTGCTCTCCCTCCGCGCGTCAATATGATGCTGCGCATATCT 120
676 AGCCAAACCAACCTGCTCTCCCTCCGCGCGTCAATATGATGCTGCGCATATCT 617
121 CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCCGCGGTTA 180
616 CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCCGCGGTTA 557
181 ATATCCGCTTATATCAAAACCGGAAGCTTATCCCGGTGCAAAAGCTTCCCTTTGCT 240
556 A-----CTAACAAACCGGAAGCTTATCCCGGTGCAAAAGCTTCCCTTTGCT 506
241 GGAAGATGCTCACTCATGTCAGACACTCTAGCTCACCGAGCAATGCTGATGAGCTGAG 300
505 GGAAGATGCTCACTCATGTCAGACACTCTAGCTCACCGAGCAATGCTGATGAGCTGAG 446
301 AATTCGAGCAAGAGGCTCATGCTTTTCACTTGAAGAGCTGCGGTGATGCTGACG 360
445 AGATACGAGCAAGAGGCTCATGCTTTTCACTTGAAGAGCTGCGGTGATGCTGACG 386
361 TGCAATCCCGAGCTGAGGAAGAGATCTGAATAGCCGCTTTTTCGATGACCGGTT 420
385 TGCAATCCCGAGCTGAGGAAGAGATCTGAATAGCCGCTTTTTCGATGACCGGTT 326
421 AAGAGATCGGCTTACTACTGATGTTTAAACAGAGCAATGCTTTTGAACACGCTGTT 480
325 AAGAGATCGGCTTACTACTGATGTTTAAACAGAGCAATGCTTTTGAACACGCTGTT 266
481 TACTGAGGAAGCTGCGCGTATGCTGCAACATCTCTTATGATCAAAACAAATCAGA 540
265 TACTGAGGAAGCTGCGCGTATGCTGCAACATCTCTTATGATCAAAACAAATCAGA 206

QY 541 AGAGCCGAGACCGACAGAGATCTCAAGCCAGATGTTGAGTTCTTGAACAC 600
 DB 205 CGGCGGAAACGACAGAGCGAGATGCTAGCCAGATGATGAGGATGCTCGAAGAACG 146
 QY 601 ---AGTAGTACGACACCTCTTTTTCGTGAGTGTCTTAAACCGCGCTTAACAC 657
 DB 145 AGCAGTACTAACCGGCTCTGTCTTTCGAGCTCTCAAGCGCGCTTACGAC 86
 QY 658 ATGATGCTCTCTATTCGACCAAGATAGACCTTGAACCAACCATGTGAGTACGT 717
 DB 85 ATGATGCTCTCTCTTTTCGACCAAGATAGACCTTGAACCAACCATGTGAGTACGT 26
 QY 718 GAAATGCTCGAAGAGTT 736
 DB 25 GAAATGCTCGAAGAGTT 7

RESULT 3
 BZ428529/c 809 bp DNA linear GSS 13-DEC-2002
 LOCUS BONHW41R BO.1.6_2_KB_tot Brassica oleracea genomic clone BONHW41,
 DEFINITION genomic survey sequence.
 ACCESSION BZ428529
 VERSION BZ428529.1 GI:26671510
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 809)
 Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 unpublished
 Other_GSSs: BONHW41R
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..809
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BONHW41"
 /clone_1lb="BO.1.6_2_KB_tot"
 /note="Vector: pROS1; Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into pROS1 using BstXI linkers"
 BASE COUNT 241 a 216 c 146 g 206 t
 ORIGIN

Query Match 31.1%; Score 495.8; DB 29; Length 809;
 Best Local Similarity 87.6%; Pred.No. 1.1e-144;
 Matches 542; Conservative 0; Mismatches 77; Indels 0; Gaps 0;

QY 975 GAGATGATATTCAGAGAGACAGACAGTGGCGCTTATCGAGAGATCCCGCTAG 1034
 DB 742 GAGATGATATTCAGAGAGACAGACAGTGGCGCTTATTCAGAGAGATCCCGCTAG 683
 QY 1035 GATGCTCTTCATCCAGATATGCAATCAACGCTACAAACAGCTGATCAAGTAGTCGG 1094
 DB 682 GATAGTGTTCATCCAGATATTCATCAACGATTCAGAGTGAACCTGATTTGATGTCGG 623
 QY 1095 GAAATCAAGAGCCCTAGATGATCTGACTTGGCTTCACTTCATATCTCAACGCGTGGT 1154
 DB 622 AAAATCAAGAGCTTAGATGATCTGACTTGGCTTCACTTCATATCTCAACGCGTGGT 563

QY 1155 GAAAGATATATGAGGCTTATCTCCAGGCCCATTCTATCATGGGCGGTTGGCAT 1214
 DB 562 AAAAGAGTCTTGAAATGATCCACAGGCCCACTTCTTATGAGGCCGCTTGGCAT 503
 QY 1215 AACAGACGATCGTTATGATGCTCTTGTTCGCGACGAGACCAACAGCAATGCTGAACAT 1274
 DB 502 AACAGATGATCTTATGATGCTCTTGTTCGCGACGAGACCAACAGCAATGCTGAACAT 443
 QY 1275 GTGGCCGCTATGATGATCCACGCTGTGGGTGATCCCTTGAAGTTAAACCTGAG 1334
 DB 442 GTGGCAATAGCGATGATCCACGCTGTGGGTGATCCCTTGAAGTTAAACCTGAG 383
 QY 1335 GTTCGTGCAAAAGAGGTGAGTGGAGTTTCCGTTCTTGAGTCCGATTTGACACTGC 1394
 DB 382 ATTCTGACGAAAGAGTGAAGTGAAGTTCGCTTCTGCTTCTGAGTTCGACTAGC 323
 QY 1395 ACCTTTCGGGTGCGGTGCTCGAATTTGCCCGGAAATCTTGTCTTACTACCGTTAT 1454
 DB 322 GCTTTTCGGGTGCTGCTGATGATCTGCTCGGAAAGATCTTGAATGACACCGTTAC 263
 QY 1455 GTTTTGAAGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1514
 DB 262 GTTCTGAGCCGAGCGCTGTTACAGATTTGAATGGGAGTGCAGATGCTAACCGGCT 203
 QY 1515 TGACTTATCTGAGAACTGAGGCTTTTTCGAGATGGCTAATCTTCTTCTGCTAATT 1574
 DB 202 TGACTTATCTGAGAACTGAGGCTTTTTCGAGATGGCTAATCTTCTTCTGCTAATT 143
 QY 1575 GCGCGCTAGCGCACTTAA 1593
 DB 142 GCGCGTAGCGCACTTAA 124

RESULT 4
 BH451907/c 674 bp DNA linear GSS 12-DEC-2001
 LOCUS BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
 DEFINITION survey sequence.
 ACCESSION BH451907
 VERSION BH451907.1 GI:17637618
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 674)
 Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 unpublished
 Other_GSSs: BOGS044TR
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..674
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOGS044"
 /clone_1lb="BOGS"
 /note="Vector: pROS1; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pROS1 using BstXI linkers"
 BASE COUNT 191 a 194 c 130 g 159 t
 ORIGIN

FEATURES
 source

Tue Jan 20 17:43:23 2004

us-10-022-025a-2.rst

Page 4

Query Match 29.9%; Score 476; DB 28; Length 674;
Best Local Similarity 89.5%; Pred. No. 1.8e-138;
Matches 512; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

QY 1022 GAGTCTCTGCTAGAGATGCTCTTCAATCCAGATATCAATCAACGCTATCAAAACAGATCG 1081
DB 674 GAGTCTCTGCTAGAGATGCTCTTCAATCCAGATATCAATCAACGCTATCAAAACAGATCG 615
QY 1082 ATCAAGATAGCGGAAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATATC 1141
DB 614 ATTCGGATAGCGGAAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCATATC 555
QY 1142 TAACGGCTGTGGTGAAGAAGATAGAGCTTCAATCCAGATATCAACGCTATCAAAACAGATCG 1201
DB 554 TGACGGCTGTGGTGAAGAAGATAGAGCTTCAATCCAGATATCAACGCTATCAAAACAGATCG 495
QY 1202 CCGGTTTGGCCATTAACAGACAGATGCTGATGCTCTTGGTTCCGACAGGACACAG 1261
DB 494 CCGGTTTGGCCATTAACAGACAGATGCTGATGCTCTTGGTTCCGACAGGACACAG 435
QY 1262 CAATGGTGAACATGTGGGCGGATGCGATGCAATGCAACAGTGGGCTTGAATCTTGGAGT 1321
DB 434 CAATGGTGAACATGTGGGCGGATGCGATGCAATGCAACAGTGGGCTTGAATCTTGGAGT 375
QY 1322 TTAACCTGAGAGGTTCTGTGGCAGAAAGATGAGGTGAGTGTTCGCTTCTTGGGTCGG 1381
DB 374 TTAACCTGAGAGGTTCTGTGGCAGAAAGATGAGGTGAGTGTTCGCTTCTTGGGTCGG 315
QY 1382 ATTGAGACTTGGACCTTTCGAGGTCGAGTGGGTCGAGATTTGGCCGAGAAAGATCTTGGT 1441
DB 314 ACTTGGACTTGGACCTTTCGAGGTCGAGTGGGTCGAGATTTGGCCGAGAAAGATCTTGGT 255
QY 1442 TTAACCTGAGAGGTTCTGTGGCAGAAAGATGAGGTGAGTGTTCGCTTCTTGGGTCGG 1501
DB 254 TGACTACCTGAGAGGTTCTGTGGCAGAAAGATGAGGTGAGTGTTCGCTTCTTGGGTCGG 195
QY 1502 ATGTAACGCGGCTTGAATCTTGAAGAACTGAGGCTTCTTGGAGATGGCTATACCTC 1561
DB 194 TTGGAACGCGGCTTGAATCTTGAAGAACTGAGGCTTCTTGGAGATGGCTATACCTC 135
QY 1562 TTCTGTGAATATGCGCCGCTAGGCGGAGTTAA 1593
DB 134 TTCTGTGTCAAGTGGCTGAGGCGGAGTTAA 103

RESULT 5
BH510149/c 745 bp DNA linear GSS 13-DEC-2001
LOCUS BOGT049TR BOGT Brassica oleracea genomic clone BOGT049, genomic
DEFINITION survey sequence.
ACCESSION BH510149
VERSION BH510149.1 GI:17718239
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryote, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 745)
1 (bases 1 to 745)
TOWN, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BOGT049TF
COMMENT TIGR Contact: Chris Town

FEATURES
1
Location/Qualifiers

source

1. 745
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000D3"
/db_xref="taxon:3712"
/clone="BOGT049"
/clone_id="BOGT"
/note="Vector: PHOS1; site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 219 a 186 c 137 g 203 t

Query Match 26.9%; Score 428; DB 28; Length 745;
Best Local Similarity 90.2%; Pred. No. 2.8e-123;
Matches 458; Conservative 0; Mismatches 50; Indels 0; Gaps 0;

QY 975 GAGATGATATTCAGAGGACAGACAGATGCGGCTTAAATGAGTGAATCCTCGTAG 1034
DB 508 GAGATGATATTCAGAGGACAGACAGATGCGGCTTAAATGAGTGAATCCTCGTAG 449
QY 1035 GATGGCTCTGATCCAGATGATGCAATCAAGGCTCAAAACAGAGTGAATCAAGTACG 1094
DB 448 GATGGCTCTGATCCAGATGATGCAATCAAGGCTCAAAACAGAGTGAATCAAGTACG 389
QY 1095 GAAATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1154
DB 388 GAAATCAAGAGCCCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 329
QY 1155 GAAAGATGATGAGGCTTCAATCCAGGCGGCTTCAATGATGAGGCGGCTTGGCCAT 1214
DB 328 GAAAGATGATGAGGCTTCAATCCAGGCGGCTTCAATGATGAGGCGGCTTGGCCAT 269
QY 1215 AACAGACAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1274
DB 268 AACAGACAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 209
QY 1275 GTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1334
DB 208 GTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 149
QY 1335 GTTCGTGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1394
DB 148 GTTCGTGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 89
QY 1395 ACCTTGGGCTGCGGCTGCGGATTTGCCCGGAGAAATCTTGGTTTACTACCGTTAT 1454
DB 88 GCCTTTGGGCTGCGGCTGCGGATTTGCCCGGAGAAATCTTGGTTTACTACCGTTAT 29
QY 1455 GTTTGACCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1482
DB 28 GTTCTGACCGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1

RESULT 6
BH527239/c 727 bp DNA linear GSS 13-DEC-2001
LOCUS BOHB006TR BOHB Brassica oleracea genomic clone BOHB006, genomic
DEFINITION survey sequence.
ACCESSION BH527239
VERSION BH527239.1 GI:17735324
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
REFERENCE Eukaryote, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 727)
1 (bases 1 to 727)
TOWN, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BOHB006TF
COMMENT TIGR Contact: Chris Town

FEATURES
1
Location/Qualifiers

Db 94 TATGCGAAGCGTGGAGGATCGCTTGAAACATCTTCACTGCTTAACAATCAAA 35

QY 541 AGAGCGAGACGACGACGAGATCTCAAGCC 574

Db 34 CGCGCGAAACGACGAGGCGAGTGAATCGCTAACC 1

RESULT 8
BZ443616 776 bp DNA linear GSS 13-DEC-2002
LOCUS BONF28TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONF28,
DEFINITION genomic survey sequence.
ACCESSION BZ443616
VERSION BZ443616.1 GI:26703871
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; rosids
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 776)
Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
REFERENCE Whole genome shotgun sequencing of Brassica oleracea
TITLE Unpublished
JOURNAL Other GSSs: BONF28TR
COMMENT Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..776
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone_lib="BO.1.6.2 KB tot"
/note="Vector: pBS1; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pBS1 using BstXI linkers"

BASE COUNT 196 a 156 c 207 g 217 t

ORIGIN

Query Match 24.6%; Score 391.8; DB 29; Length 776;
Best Local Similarity 81.1%; Pred. No. 7,8e-112;
Matches 481; Conservative 0; Mismatches 107; Indels 5; Gaps 2;

QY 967 GTCTTTGGAGATGATATTCAGAGAACAGACAGAGTGGCGCTTATCGAGGATC 1026

Db 184 GTTTTACAGGAATGATATTAGAGAACTGACACGGTGGCTGATTCAGAGGAT 243

QY 1027 CTGGCTTGAATGCTCTTCACTCAAGTATGCAATCAACGGTACAAAGAGTGGATTA 1086

Db 244 CTGGCTTGAATGCTCTTCACTCAAGTATGCAAGTATGCAAGTATGCAAGTATGCAAG 303

QY 1087 GTAGTGGGAAATCAAGAGCCCTAGATGATCTGACTTGGCTTCACTTCAATCTAACG 1146

Db 304 GTGTGAGGAGATCAAGAGCCCTAGATGATCTGACTTGGCTTCAATCTAACG 363

QY 1147 GCTGTGTGAAGAATATTTAGAGGCTTCACTTCCAGAGCCCTTCAATCTAACG 1206

Db 364 GCTGTGTGAAGAATATTTAGAGGCTTCACTTCCAGAGCCCTTCAATCTAACG 423

QY 1207 TTGGCCATTAACAGACGATCGTTGATGCTGCTTCTTCCGCAAGGAGCAACGATG 1266

Db 424 TTAGCAATCAACAGATGATCTTGAAGGCGGCTGCTGCGGCAAGGAGCACTGCAATG 483

QY 1267 GTAAATGTGGGCGGATGATGATCAACAGTGTGGCTTGAATCTTTGAGTTTAA 1326

Db 484 GTTAACATGTGGCTATGACCCATGATCCACACGTGGGAAATCGCTTGAGTTGAA 543

QY 1327 CCTGAGAGTTTGTGGCAAAAGAGTGAAGTGAATTTTGGTTCTTGGGCGATTTG 1386

Db 544 CCGGAAGGTTTGTGGCAAAAGAGTGAAGTGAATTTTGGTTCTTGGGCGATTTG 603

QY 1387 AGACTTGAACCTTTGGGCTGGGCTGCTGCAATTTTCCCGGGAAGATCTTGGTTTACT 1446

Db 604 AGCTTGAACCTTTGGGCTGGGCTGCTGCAATTTTCCCGGGAAGATCTTGGTTTACT 663

QY 1447 ACCGTTATGTTTGAACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 1503

Db 664 ACCGTTATGTTTGAACCGGATGATGATGATGATGATGATGATGATGATGATGATGAT 723

QY 1504 GGTAAACG--CGTGAATTAATCTGAAGAACTGAGGCTTTCTTGGGATGCT 1554

Db 724 GATGACAGACCGTGAATCTTGTCCAGAAACTGAGGCTTCTGTGAGATGCT 776

RESULT 9
BH977803 705 bp DNA linear GSS 02-OCT-2002
LOCUS odel4a11.b1 B.oleracea002 Brassica oleracea genomic, genomic survey
DEFINITION sequence.
ACCESSION BH977803
VERSION BH977803.1 GI:23460806
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Eukaryota; Magnoliophyta; eudicotyledons; core eudicots; rosids
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 705)
Delehaunty,K., Fawell,G., Fulton,L., McCombie,W.R., Miner,T., Nash
M., Rabinowicz,P.D. and Wilson,R.K.
REFERENCE Whole genome shotgun reads from Brassica oleracea
TITLE Unpublished
JOURNAL Contact: Richard K. Wilson
COMMENT Genome Sequencing Center
Washington University School of Medicine
Email: submissions@wustl.wustl.edu
Plate: odel4 row: a column: 11
Seq primer: -21UPPOT forward
Class: shotgun
High quality sequence start: 18
High quality sequence stop: 519.
Location/Qualifiers
1..705
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/db_xref="taxon:3712"
/clone_lib="B.oleracea002"
/note="Vector: pCRW13; Whole genome shotgun library from
flowering buds. DNA was purified from a crude nuclear
Thomas Osborn at the University of Wisconsin. Genomic
DNA was provided by Fabio Rabinowicz (CSHL) and the
shotgun library prepared at Washington University Genome
Sequencing Center."

BASE COUNT 173 a 143 c 193 g 193 t 3 others

ORIGIN

Query Match 20.8%; Score 330.8; DB 28; Length 705;
Best Local Similarity 76.1%; Pred. No. 1.4e-92;
Matches 458; Conservative 0; Mismatches 139; Indels 5; Gaps 4;

QY 967 GTCTTTGGAGATGATATTCAGAGAACAGACAGAGTGGCGCTTATCGAGGATC 1026

Db 103 GTTTTACAGGAATGATATTAGAGAACTGACACGGTGGCTGATTCAGAGGAT 162

QY 1027 CTGGCTTGAATGCTCTTCACTCAAGTATGCAATCAACGGTACAAAGAGTGGATTA 1086

Db 163 CTTCGTATGATGCTCTTCATCCAGATATTCAGTCAATGCTTCAACAAGCTTGATC 222

Qy 1087 GTAGTCGGGAAATCAAGAGCCCTAGATGATCTGATCTGCTTCACTTCATATCTAACG 1146

Db 223 AACGTGGGAAGATCAAGAGCCCTAGATGATCTGATCTGCTTCACTTCATATCTGACT 282

Qy 1147 GCTGTGATGAAGAATGATTTAGAGCT-TCATCTCCAGGCCCCAG-TTCATATGATGGGCC 1204

Db 283 GCTGTGTGAAAGATGCTTGAAGCTTCAACCGCAGGCCCACTTCTGTGTGGGCC 342

Qy 1205 GTTGGCCATTAAGACAGATGCTGATGCTGCTTCTTCCGAGGAGACACAGCAA 1264

Db 343 GCTTAGCATCAAGACATATTAACCGCCGTGTGATCCGGAAGGTACACCGCAA 402

Qy 1265 TGGTGAACATGTGGCCGATCGCATGATCCACAGTGTGGTTGATCTTTGAGTTTA 1324

Db 403 TGGTGAACATGTGGCTATAGCAATGATCCACAGTGTGGGAAATCTTTGAGTTTA 462

Qy 1325 AACCTGAGAGTTCGTGGCAAAAGAGTGAAGTGAAGTTTCGTTCTTGGGTCGATT 1384

Db 463 AACCTGACGCTTGTGAGCAAGAGTGAAGGAGGAGTTCTGGTTCTTGGATCGGATC 522

Qy 1385 TGAAGCTTGAACCTTTCGGGTCGGTCTGCGATTTGCCCGGAGAAATCTTGTGTTTA 1444

Db 523 TGAAGCTGACCTTTCGGGCGGGGCGGTGGGCTGCTCTGAAAGAAATCTTGAAGTTAG 582

Qy 1445 CTACCGTATGTTTGAAGCGCGATGATGTTACATGATGTTGAATG--GGAACGTCGA 1502

Db 583 CACGCTATGATTTGAAGCGCGATCTCTTCCCGAGTGAAGTGTTCACCTTACAG 642

Qy 1503 TGAAGCGCGCTTGAATCTGAGAAA-CTGAGGCTTCTTTCGAGATGCTTATCTTC 1561

Db 643 TGAAGAGCTTGAATCTTTCGAGAAATGAGGCTTTCGATGAGATGGCTTATCTTC 702

Qy 1562 TT 1563

Db 703 TT 704

RESULT 10

BH534160 732 bp DNA linear GSS 14-DEC-2001

LOCUS BOHC239TF BOHC Brassica oleracea genomic clone BOHC239, genomic

DEFINITION survey sequence.

ACCESSION BH534160

VERSION BH534160.1 GI:17761104

KEYWORDS GSS.

SOURCE Brassica oleracea

ORGANISM Brassica oleracea

REFERENCE Town C.D., Van Aken S., Uetreback T., Koo H. and Fraser C.M.

AUTHORS Whole genome shotgun sequencing of Brassica oleracea

TITLE Unpublished

JOURNAL Other GSSs: BOHC239TF

COMMENT Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtowne@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TF

Class: sheared ends.

Location/Qualifiers

1..732

location="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO100DH3"

/db_xref="taxon:3712"

/clone="BOHC239"

BASE COUNT 196 a 164 c 169 g 203 t

ORIGIN

Query Match 20.8%; Score 330.8; DB 28; Length 732;

Best Local Similarity 75.2%; Pred. No. 1.4e-92;

Matches 445; Conservative 0; Mismatches 132; Indels 15; Gaps 2;

238 GTTGAAGCATGTCATCATGTCAGACATCTTACTTACCGACGACATCGGTGATGAGCT 297

Db 141 GTCCGAGACATGTCATCATGTCATCCAGCTTTTACTACCGCTGATAGCGACGCCG 200

Qy 298 GAGAAATTCGAGGCAAGAGGCTCATGCTTTCAGTTAGAGAGACTCGGTGATGCTC 357

Db 201 GAGAAATTCGAGGCAAGAGGCTTTCAGTTAGAGAGACTCGGTGATGCTC 260

Qy 358 ACGTGAATTCGAGGCAAGAGGCTTTCAGTTAGAGAGACTCGGTGATGCTC 417

Db 261 ACGTGAATTCGAGGCAAGAGGCTTTCAGTTAGAGAGACTCGGTGATGCTC 320

Qy 418 GTTAAAGATGCGCTTACTCATGATGTTTAAAGAGAGATTTGGTTTGCACACAGCGT 477

Db 321 GTTAAAGATGCGCTTACTCATGATGTTTAAAGAGAGATTTGGTTTGCACACAGCGT 380

Qy 478 GTTAACTGCGCAACGCTTTCGCGATTCGCTTGAACCATCTCTTGTATCAAAACAATC 537

Db 381 GTTAACTGCGCAACGCTTTCGCGATTCGCTTGAACCATCTCTTGTATCAAAACAATC 440

Qy 538 AGAAGAGCGGAGAGCGGCAAGAGGATGATCAAGCGAGATGTTGATTTCTTGAAGAA 597

Db 441 AAGCGCTTGAAGAGCGGCAAGAGGATGATCAAGCGAGATGTTGATTTCTTGAAGAA 500

Qy 598 CAGAGTGAAGCA-----ACCTGTTTGTTCGTGATGTTGCTTAAAGCGCGTCTT 651

Db 501 CAGAGCGGCAAGAGGATGATGTTTGTTCGTGATGTTGCTTAAAGCGCGTCTT 560

Qy 652 AACAATATATGCTCTGATATTCGACCAAGATGATGATGATGATGATGATGATGATG 706

Db 561 AATAATATATGCTCTGATATTCGACCAAGATGATGATGATGATGATGATGATGATG 620

Qy 707 ----TTGATTAAGTGAATGCTGCAAGAGGTTATGATTTGCTCGAAGCTTGAATGG 762

Db 621 GTGATGAGCTTCGTGCTTGTGTCGAAGAGGTTATGATTTACTCGAAGCTTGAATGG 680

Qy 763 ACTGATACCTTCTTGTGCTATGCAAGATTTGATCTTCAAGACTCCGCTTA 814

Db 681 ACCGATCATCTCCGCTGCTGCAAGATTTGATCTTCAAGACTCCGCTTA 732

RESULT 11

CB344213 826 bp mRNA linear EST 10-APR-2003

LOCUS CA48EN001.1af G02 Cabernet Sauvignon Leaf-CA48EN Vitis vinifera

DEFINITION cDNA clone CA48EN001.1af_G02 5', mRNA sequence.

ACCESSION CB344213

VERSION CB344213.2 GI:29784005

KEYWORDS EST.

SOURCE Vitis vinifera

ORGANISM Vitis vinifera

REFERENCE Goes da Silva, F., Lim, H., Tandolino, A., Baek, J., Jones, K., Walker

AUTHORS M.A. and Cook, D.R.

TITLE Transcriptional responses of Vitis vinifera to infection by the

JOURNAL bacterial pathogen Xylella fastidiosa

COMMENT On Mar 14, 2003 this sequence version replaced gi:28965180.

Contact: Douglas Cook, PhD

CBS Genome Facility

UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGACATATGCC.
Location/Qualifiers
1. 826

FEATURES

source
/organism="Vitis vinifera"
/mol_type="rRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA48EN001_1bf_902"
/sex="Hermaphrodite"
/dev_stage="Berry stage I"
/lab_host="DH5alpha"
/note="Organ: Berry; Vector: pGEMT-1; Site: 1; Site 2: 5'-ATCTGAGCTTGGTCACTTCAATCTAAAGGCTGTGGTAAAGATATTTAGAGGCTC 1114
1183 AATCCGACATTCGGGAGATGTAACCTGAGGCGGTCTGTAAGAGGTTCTGAGGCTGC 242
1175 ATCTCGAGGCCACTTCTATCAATGAGGCGGTGGCCATTAACAGACGATCGGTGATG 1234
243 ACCACCGGCGCCACGCTCTCGTGGCGCCGTTTATCAATCAACATTCACACGCGTGA 302
1235 GTCGTCTTGTTCGGGAGGAGACACAGCAATGATGATGAGGCGGTATCGATGATC 1234
303 GCATACGCTGCTTCGCGGTCCACAGCATGTTTAACTGTGGGCTATCAACGAGATC 362
1295 CACACGTGTGGTGAATCTTTGAGATTAACTGAGAGGTTGCTGGCAAAAGAGTG 1354
363 CACAGATGTGTGATGATCACTGACCTGACGCGGACGATTCGTGACACCCCTGCTG 422
1355 AGGTGAGATTTGCTGCTTGGGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 1414
423 ACGTGAATCTCTCTGTTTGAATCGATCTCAAGGTCGCGCATTTGGGTCGAGCAAA 482
1415 GATTTGGCCCGGAGAAATCTTGTGTTTAACTGATGATGTTTGAACGGGATGATGT 1474
483 GAACCTGCCCCGGAAGAACCTTGTGCTGACACAGCCACTTTTGGGTGGCGTGGCTGT 542

BASE COUNT

179 a 218 c 224 g 203 t

ORIGIN

Query Match 20.7%; Score 329.4; DB 14; Length 826;
Best Local Similarity 68.7%; Pred. No. 4,3e-92;
Matches 433; Conservative 0; Mismatches 206; Indels 0; Gaps 0;

935 CAGATTAATTTATCCGACCCGACATATATCGCGTCTTTGGAGATGATATTCAGAGAA 994
3 CCATTAAGTTGTGATTCGATATATCGCGTCTTTGGAGATGATATTCAGAGAA 62
995 CAGACAGATTTGGCTTAATCGAGTGAATCGTCTGATGATGATGATGATGATGATGAT 1054
63 CGATTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 122
1055 TGAATCAACGATCAAAACGAGCTGATCAAGTATGATGATGATGATGATGATGATGAT 1114
123 TCCATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 182
1115 AATCTGAGCTTGGTCACTTCAATCTAAAGGCTGTGGTAAAGATATTTAGAGGCTC 1114
183 AATCCGACATTCGGGAGATGTAACCTGAGGCGGTCTGTAAGAGGTTCTGAGGCTGC 242
1175 ATCTCGAGGCCACTTCTATCAATGAGGCGGTGGCCATTAACAGACGATCGGTGATG 1234
243 ACCACCGGCGCCACGCTCTCGTGGCGCCGTTTATCAATCAACATTCACACGCGTGA 302
1235 GTCGTCTTGTTCGGGAGGAGACACAGCAATGATGATGAGGCGGTATCGATGATC 1234
303 GCATACGCTGCTTCGCGGTCCACAGCATGTTTAACTGTGGGCTATCAACGAGATC 362
1295 CACACGTGTGGTGAATCTTTGAGATTAACTGAGAGGTTGCTGGCAAAAGAGTG 1354
363 CACAGATGTGTGATGATCACTGACCTGACGCGGACGATTCGTGACACCCCTGCTG 422
1355 AGGTGAGATTTGCTGCTTGGGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 1414
423 ACGTGAATCTCTCTGTTTGAATCGATCTCAAGGTCGCGCATTTGGGTCGAGCAAA 482
1415 GATTTGGCCCGGAGAAATCTTGTGTTTAACTGATGATGTTTGAACGGGATGATGT 1474
483 GAACCTGCCCCGGAAGAACCTTGTGCTGACACAGCCACTTTTGGGTGGCGTGGCTGT 542

QY 1475 TACATGATTTGATGAGGAGACCTCCGATGATAGCGGCTTGAATTCAGAAATCTGA 1534
DB 543 TGCATGAGTGTGAGGAGGCTTCATCCGACCCGAGGATCTATCCGAGGTGTGA 602
QY 1535 GGCCTTCTTGGGAGATGAGCTATATCTTCTGCTGATTAATTTGCGCGGAGGATTTAA 1593
DB 603 GGCATCTCTGCGAATGAGCACACCTCTGACTGTGAGGAGTGGGCGCCAGGCAATGA 661

RESULT 12
CB343993 638 bp mRNA linear EST 10-APR-2003
LOCUS CA48EN001_1bf_902 Cabernet Sauvignon leaf-CA48EN Vitis vinifera
DEFINITION CDNA clone CA48EN001_1bf_902 5', mRNA sequence.
CB343993
ACCESSION CB343993.1 GI:28964960
VERSION EST
KEYWORDS
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
Vitis vinifera
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; Vitaceae; Vitis.
1 (bases 1 to 638)
Goes de Silva, P., Lim, H., Iandolo, A., Baek, J., Jones, K., Walker
M.A. and Cook, D.R.
Transcriptional responses of Vitis vinifera to infection by the
bacterial pathogen Xylella fastidiosa
Unpublished
Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: ACGGTACCGACATATGCC.
Location/Qualifiers
1. 638

FEATURES

source
/organism="Vitis vinifera"
/mol_type="rRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CA48EN001_1bf_902"
/sex="Hermaphrodite"
/dev_stage="Berry stage I"
/lab_host="DH5alpha"
/note="Organ: Berry; Vector: pGEMT-1; Site: 1; Site 2: 5'-ATCTGAGCTTGGTCACTTCAATCTAAAGGCTGTGGTAAAGATATTTAGAGGCTC 1114
1183 AATCCGACATTCGGGAGATGTAACCTGAGGCGGTCTGTAAGAGGTTCTGAGGCTGC 242
1175 ATCTCGAGGCCACTTCTATCAATGAGGCGGTGGCCATTAACAGACGATCGGTGATG 1234
243 ACCACCGGCGCCACGCTCTCGTGGCGCCGTTTATCAATCAACATTCACACGCGTGA 302
1235 GTCGTCTTGTTCGGGAGGAGACACAGCAATGATGATGAGGCGGTATCGATGATC 1234
303 GCATACGCTGCTTCGCGGTCCACAGCATGTTTAACTGTGGGCTATCAACGAGATC 362
1295 CACACGTGTGGTGAATCTTTGAGATTAACTGAGAGGTTGCTGGCAAAAGAGTG 1354
363 CACAGATGTGTGATGATCACTGACCTGACGCGGACGATTCGTGACACCCCTGCTG 422
1355 AGGTGAGATTTGCTGCTTGGGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAG 1414
423 ACGTGAATCTCTCTGTTTGAATCGATCTCAAGGTCGCGCATTTGGGTCGAGCAAA 482
1415 GATTTGGCCCGGAGAAATCTTGTGTTTAACTGATGATGTTTGAACGGGATGATGT 1474
483 GAACCTGCCCCGGAAGAACCTTGTGCTGACACAGCCACTTTTGGGTGGCGTGGCTGT 542

BASE COUNT

133 a 172 c 189 g 144 t

ORIGIN

Query Match 20.3%; Score 322.6; DB 14; Length 638;
Best Local Similarity 69.4%; Pred. No. 5.1e-90;
Matches 439; Conservative 0; Mismatches 154; Indels 0; Gaps 0;

935 CAGATTAATTTATCCGACCCGACATATATCGCGTCTTTGGAGATGATATTCAGAGAA 994
3 CCATTAAGTTGTGATTCGATATATCGCGTCTTTGGAGATGATATTCAGAGAA 62

QY 995 CAGACAGATTGCGGCTTAAATCGAGTGAATCTCGTAGAGTGGCTTCATCCAGATA 1054
 Db 63 CGGATACGGTGGCAGTTTGAATGAGTGGATCTCGACAGGATGGTACTCCACCTGAGT 122
 QY 1055 TCCATCAACGGTACAAAACGAGTGAATCAAGTATGTCGGAATCAAGACCTTATG 1114
 Db 123 TCCATCAACGGTACAAAACGAGTGAATCAAGTATGTCGGAATCAAGACCTTATG 182
 QY 1115 AATCTGACTTGGCTTCACTTCCATATCAAGCGGCTGATGAAGAATGATGAGCTTC 1174
 Db 183 AATCCGACATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 242
 QY 1175 AATCCGACATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1234
 Db 243 AATCCGACATTCGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 302
 QY 1235 GTGCTTCTTCTCGGACAGGACACAGCATGATGATGATGATGATGATGATGATGATG 1294
 Db 303 GTGCTTCTTCTCGGACAGGACACAGCATGATGATGATGATGATGATGATGATGATG 362
 QY 1295 CACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1354
 Db 363 CACAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 422
 QY 1355 AGGTGAGTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1414
 Db 423 AGGTGAGTTCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 482
 QY 1415 GATTTGCGCGGAGGAGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1474
 Db 483 GATTTGCGCGGAGGAGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 542
 QY 1475 TACATGAGTTGATGAGGAGGAGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 1534
 Db 543 TACATGAGTTGATGAGGAGGAGATCTGCTTCTGCTGCTGCTGCTGCTGCTGCTG 602
 QY 1535 GGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1567
 Db 603 GGTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 635

RESULT 13
 BG648529
 LOCUS EST10148 HOGA Medicago truncatula cDNA clone phOGA-2355 5' end,
 DEFINITION mRNA sequence.
 ACCESSION BG648529
 VERSION BG648529.1 GI:13783641
 KEYWORDS EST.
 SOURCE Medicago truncatula (barrel medic)
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eurosid1; Fabales; Fabaceae; Papilionoideae; Trifoliales;
 Medicago.
 1 (bases 1 to 720)
 Hahn,M.G., Ojane-Reuhs,T., Samac,D., Town,C.D., Van Aken,S.,
 Uteback,T., Cho,J. and Fraser,C.M.
 ESTs from roots of Medicago truncatula treated with
 oligogalacturonides of DP 6-20
 Unpublished
 Contact: Michael G. Hahn
 Complex Carbohydrate Research Center
 University of Georgia
 220 Riverbend Road, Athens, GA 30602-4712, USA
 Tel: 706-542-4457
 Fax: 706-542-4412
 Email: hahn@ccrc.uga.edu
 G392133e TIGR sequence name: MTMCM27TK More information is
 available at: www.medicago.org
 Seq primer: Skmod (CTA GAA CTA gta gat CC).
 Location/Qualifiers
 1..720

BASE COUNT 179 a 128 c 201 g 212 t

ORIGIN

Query Match 19.8%; Score 314.8; DB 10; Length 720;
 Best Local Similarity 66.7%; Pred. No. 1,6e-87;
 Matches 483; Conservative 0; Mismatches 232; Indels 9; Gaps 2;

716 GTGAATGCTGGAAGAAGTTATGATGATGCTGCGAAAGTTGATGATGATGATGATGATG 775
 Db 2 GTGAGTGTGTTGAAACAAGTTATGATGATGATGATGATGATGATGATGATGATGATG 61

776 CTGCGATATCGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 835
 Db 62 CATTTGAAAGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 121

836 AGTTAAACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 895
 Db 122 AAGTAAACCGGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181

896 CTCGTGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 955
 Db 182 ACAAGATTTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241

956 ACATTAATGCGCGTCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1015
 Db 242 ACATTAATGCGCGTCTTGGGAGATGATGATGATGATGATGATGATGATGATGATGATG 301

1016 TCGATGATCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1075
 Db 302 TAGATGATCTCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 361

1076 AGCTGATCAAGT---AGTCGGGAAATCAAGACCCCTGATGATGATGATGATGATGATG 1132
 Db 362 AGCTGATCAAGT---AGTCGGGAAATCAAGACCCCTGATGATGATGATGATGATGATG 421

1133 TTCAATATCAACGGTACAAAACGAGTGAATCAAGTATGTCGGAATCAAGACCTTATG 1192
 Db 422 TTCAATATCAACGGTACAAAACGAGTGAATCAAGTATGTCGGAATCAAGACCTTATG 481

1193 TATCATGAGGCGGCTTGGCCATTAACAGACAGATGATGATGATGATGATGATGATGATG 1252
 Db 482 TATCATGAGGCGGCTTGGCCATTAACAGACAGATGATGATGATGATGATGATGATGATG 541

1253 GATCCACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1312
 Db 542 GATCCACAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 601

1313 CTTGGATTTAAACCTGAGAGGTTCTGCGAAAGAGGTGAGGTGAGGTGATGATGATG 1372
 Db 602 CACTTGAATTAATCCGAGAGGTTCTGCGAAAGAGGTGAGGTGAGGTGATGATGATGATG 655

1373 TTGGGTGATTTAGACTTGACCTTGGGCTGCGGCTGCGGATTTGCGGAGGAAG 1432
 Db 656 TTGGGTGATTTAGACTTGACCTTGGGCTGCGGCTGCGGATTTGCGGAGGAAG 715

QY 1433 ATCT 1436
 DB 716 ATTT 719
 RESULT 14
 LOCUS BH451899 576 bp DNA linear GSS 12-DEC-2001
 DEFINITION BOGS044TF BOGS Brassica oleracea genomic clone BOGS044, genomic survey sequence.
 ACCESSION BH451899
 VERSION BH451899.1 GI:17637610
 KEYWORDS GSS.
 ORGANISM Brassica oleracea
 SOURCE Brassica oleracea
 REFERENCE 1 (bases 1 to 576)
 AUTHORS Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished
 COMMENT Other GSSs: BOGS044TR
 CONTACT: Chris Town
 TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.
 FEATURES
 source Location/Qualifiers
 1..576
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOGS044"
 /clone_1ib="BOGS"
 /note="Vector: pROSL, Site 1, BstXI, 2-3 kb sheared genomic DNA inserted into pROSL using BstXI linkers"
 BASE COUNT 137 a 162 c 114 g 163 t
 ORIGIN
 Query Match 18.8%; Score 299.6; DB 28; Length 576;
 Best Local Similarity 83.9%; Pred. No. 8.7e-83;
 Matches 354; Conservative 0; Mismatches 59; Indels 9; Gaps 1;
 QY 1 ATGGCTACGAAAGCTCGAAGCTCTTAATCTTGGCCCTTTTGTCCAAATGACGCGTTCTA 60
 DB 164 ATGGCTACGAAAGCTCGAAGCTCTTAATCTTGGCCCTTTTGTCCAAATGACGCGTTCTA 223
 QY 61 AGCCAAACCAACCTTGGCTTCTCCCTCGCCGCTGACATCATCTGCTCCGCAATCT 120
 DB 224 AGCCAAACCAACCTTGGCTTCTCCCTCGCCGCTGACATCATCTGCTCCGCTTCT 283
 QY 121 CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATATCTTCCGCGGTTA 180
 DB 284 CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATATCTTCCGCGGTTA 343
 QY 181 ATATCCGCTTATCAAAACCGGAAAGCTTATCCCGGTGACCAAAAGCTTCCCTTGGTT 240
 DB 344 A-----CTAAACAAACCGGAAAGCTTATCCCGGTGACCAAAAGCTTCCCTTGGTT 394
 QY 241 GGAAGCATCTCACTCATGTCAAGCACTTATGCTCAGCAGGAATGCTATGACAGTGA 300
 DB 395 GGAAGCATCTCACTCATGTCAAGCAGCTTATGCTCAGCAGGAATGCTATGACAGTGA 454
 QY 301 AATTCGAGGACCAAGAGCTCATGCTTCAAGCTTGAAGAGAGCTGCGGTATGCTCAG 360
 DB 455 AGATACGAGGACCAAGAGCTCATGCTTCAAGCTTGAAGAGAGCTGCGGTATGCTCAG 514

QY 361 TCCATCCCGACGCTAGCAAGAGATCTGAATAGCCCGGTTTGTGATGACGCGGTT 420
 DB 515 TCCATCCCGACGCTAGCAAGAGATCTGAATAGCCCGGTTTGTGATGACGCGGTT 574
 QY 421 AA 422
 DB 575 AA 576
 RESULT 15
 LOCUS BG839564/c 791 bp mRNA linear EST 25-MAY-2001
 DEFINITION BG839564 F Gm01 AAFc ECORC Glycine max cold stressed_leaves
 ACCESSION BG839564
 VERSION BG839564.1 GI:14205866
 KEYWORDS EST.
 ORGANISM Glycine max (soybean)
 SOURCE Glycine max
 REFERENCE 1 (bases 1 to 791)
 AUTHORS Singh, J.A., Robert, L.S., Lu, B., Zhu, L., De Moore, A., Couroux, P., Harris, L.J., Hattori, J.I., Ouellet, T., Sprout, D. and Tinker, N.A.
 TITLE Expressed Sequence Tags from Cold-Stressed Glycine max (soybean) leaves
 JOURNAL Unpublished
 COMMENT Contact: Singh, J.A.
 Eastern Cereal and Oilseed Research Centre
 Agriculture and Agri-Food Canada
 KM Neatby Bldg., Central Experimental Farm, Ottawa, Ontario, K1A 0C6, Canada
 Tel: (613) 759-1662
 Fax: (613) 759-1701
 Email: singhja@agr.ca
 Location/Qualifiers
 1..791
 /organism="Glycine max"
 /mol_type="mRNA"
 /cultiivar="Maple Arrow"
 /db_xref="taxon:3847"
 /clone="Gm01_13h01"
 /tissue_type="leaves"
 /dev_stage="15-day seedlings"
 /clone_1ib="Gm01_AAFc_ECORC_Glycine_max_cold_stressed_leaves"
 /note="Vector: Bluescript SK-/XhoI-ECORI, Site 1; ECORI, Site 2; XhoI; Plants were grown 12 days from seeds, treated at 20C for 3 days, then leaves were collected. Library was prepared using Unizap ZAP-CDNA synthesis kit / Packaged GigaPack III Gold."
 BASE COUNT 202 a 243 c 179 g 166 t 1 others
 ORIGIN
 Query Match 18.7%; Score 298.6; DB 12; Length 791;
 Best Local Similarity 69.3%; Pred. No. 2.2e-82;
 Matches 481; Conservative 0; Mismatches 200; Indels 13; Gaps 5;
 QY 896 CTGCTGATTTGCTGACGCTTTTGTCT--CTCCCTTCATGCTTCAATTAATTAATCCGAC-C 952
 DB 758 CCGCGATTTGCTGACGCTTTTGTCTCTCTTCCAGAGTCCGATTAATTTGCTTCACTT 699
 QY 953 CGGACATTAATGCGCTTCTTTGGAGATGATATTCAGAGAACACACAGTT-GCGGTC 1011
 DB 698 CCGACATTAATGCGCTTCTTTGGAGATGATATTTAGGGGACCGACACCGGTGCGGTT 639
 QY 1012 TTAATCGAGTGAATCTCTGCTAGAGTGTCTTATCAATATGCAATCAAGCTTCAA 1071
 DB 638 TTGATTTGATGATATCTGCGAGAGATGATGATGATCGGAGGTGCAAAAGAGTCCAA 579
 QY 1072 AAGAGCTGATCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1131

```
Db 578 GAGGAGTTGACCGCGGTGTTAGG---GGTGGCGCTTTGACGAGAGGTCTGTGGCGCG 522
QY 1132 CTTCATATCTAACGGCTGTGTGAAAGAGTATGAGGCTTCATCTCCAGGCCACTT 1191
Db 521 ACGCGATCTTGTGGGGGTGTGAAGAGTTCTGAGGCTGCACCGCGGGCCGCTT 462
QY 1192 CTATCATGGGCCCCGTTGGCCATPACAGACACGATGCTGATGTGTCTTGTCCGGCA 1251
Db 461 CTCTGGGGCCCGCTGGCCATCACTGATPACAGACATGATGGTATCACGTCCTGCG 402
QY 1252 GGGACCAAGCAATGTGAACATGTGGCCGTATCCATGATCCACACGTGTGGTTGAT 1311
Db 401 GGGACCAAGCAATGTGAACATGTGGCCGTATCCATGATCCACACGTGTGGTTGAT 1311
QY 1312 CCTTGGAGTTTAACTGTGAGGTTGTGTGCAAAAGAGTGAAGTGTGCTT 1371
Db 341 CCACTTGAGTTCAAGCCGAGAGGTTCATGGGTCTGGAA-----ACGAGTTTCTGTT 288
QY 1372 CTGGGTCGGATTGAGACTTGACCTTTGGGTCGGGTGTGTGCAATTTGCCCGGGAG 1431
Db 287 TTGGGTCGGATCTGAGACTCGCTCCATTTGGGTCGGGTGTGTGCAATTTGCCCGGGAG 228
QY 1432 AATCTGGTTTAACTGACGTTATGTGTTGACGCGCAGTATGTTACATGAGTTGAATG 1491
Db 227 ACTTTGGGTTGAGCACCGTAACTCTGGGTGGCTTGGCTTTGCATGAGTTGAATG 168
QY 1492 GGAACCGTCGATGTGTAAGCGCGTTGACTTATCTGAGAACTGAGGCTTCTTGGAGATG 1551
Db 167 CTACCGTCTGATGAGCCAAAGTTGATCTAACGAGGTGCTGAGGCTCTGTGTGAATG 108
QY 1552 GCTAATCTCTTCTGCTAAATGCGCGGTAGGC 1585
Db 107 GCTAATCTCTTCTGCTAAATGCGCGGTAGGC 74
```

Search completed: January 17, 2004, 04:15:21
Job time : 3524 secs

RESULT 2

BZ380060
LOCUS BZ380060 262 bp DNA linear GSS 26-NOV-2002
DEFINITION SALK_114536.29.00.x Arabidopsis thaliana TDNA insertion lines
Arabidopsis thaliana genomic clone SALK_114536.29.00.x, genomic
survey sequence.
BZ380060
ACCESSION BZ380060.1 GI:25472537
VERSION
KEYWORDS
SOURCE GSS
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Arabidopsi.
1 (bases 1 to 262)
Alonso, J.M., Leisner, T.J., Barajas, P., Chen, H., Cheuk, R., Gadgil, N.
, C., Jeske, A., Kames, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P.,
, Zimmermann, J., and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the
Arabidopsis Genome
Unpublished
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At2g46660.
Class: TDNA tagged.
Location/Qualifiers
1..262
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_114536.29.00.x"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html"
BASE COUNT 64 a 56 c 62 g 75 t 5 others
ORIGIN
Query Match 3.1%; Score 50; DB 29; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.3e-16;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1112 ATGATCTGACTGCTGCTCACTTCATCTAACGGCTGTGGTGAAGA 1161
Db 160 ATGATCTGACTGCTGCTCACTTCATCTAACGGCTGTGGTGAAGA 209

RESULT 3
BZ451907/c
LOCUS BZ451907 674 bp DNA linear GSS 12-DEC-2001
DEFINITION BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
survey sequence.
BZ451907
ACCESSION BZ451907.1 GI:17637618
VERSION
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 674)
Town, C.D., Van Aken, S., Uteback, T., Koo, H., and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea

COMMENT
Other_GSSs: BOGS044TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..674
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGS044"
/clone="BOGS044"
/note="Vector: pROS1, site 1: BstXI, 2-3 kb sheared
genomic DNA inserted into pROS1 using BstXI linkers"
BASE COUNT 191 a 194 c 130 g 159 t
ORIGIN
Query Match 2.7%; Score 43; DB 28; Length 674;
Best Local Similarity 100.0%; Pred. No. 1.7e-12;
Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1286 CGCATGATCCACACGTGTGATGCTTTGAGTTAAACC 1328
Db 410 CGCATGATCCACACGTGTGATGCTTTGAGTTAAACC 368

RESULT 4
BH510149/c
LOCUS BH510149 745 bp DNA linear GSS 13-DEC-2001
DEFINITION BOGT049TR BOGT Brassica oleracea genomic clone BOGT049, genomic
survey sequence.
BH510149
ACCESSION BH510149.1 GI:17718239
VERSION
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosid II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 745)
Town, C.D., Van Aken, S., Uteback, T., Koo, H., and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Other_GSSs: BOGT049TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
Location/Qualifiers
1..745
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOGT049"
/clone="BOGT049"
/note="Vector: pROS1, site 1: BstXI, 2-3 kb sheared
genomic DNA inserted into pROS1 using BstXI linkers"
BASE COUNT 219 a 186 c 137 g 203 t
ORIGIN
Query Match 2.7%; Score 43; DB 28; Length 745;

Page 3

QY 996 AGACACAGTTGCGGTCTTATCGAGTGCATCTCGTA 1033
|||||
DB 45 AGACACAGTTGCGGTCTTATCGAGTGCATCTCGTA 82

USA
7912 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208

Tue Jan 20 17:43:21 2004

us-10-022-025a-2.011go.rst

Page 4

Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES

source

Location/Qualifiers

1..712

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOHRW74"

/clone_id="BO_1.6_2_KB_tot"

/note="Vector: PHOS1; Site_1: BstXI; 1.6-2 kb sheared

total DNA inserted into PHOS1 using BstXI linkers"

168 a 179 c 177 g 188 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 2.2%; Score 35; DB 29; Length 712;

Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

637 AAACGGCGGCTTAAACATGATGCTCTGT 671

DB

387 AAACGGCGGCTTAAACATGATGCTCTGT 421

RESULT 8

LOCUS

BOHRT46TR BOHR Brassica oleracea genomic clone BOHRT46, genomic

DEFINITION

survey sequence.

ACCESSION

U000000000

VERSION

U000000000

KEYWORDS

SSS.

SOURCE

Brassica oleracea

ORGANISM

Brassica oleracea

REFERENCE

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

AUTHORS

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

TITLE

; eurosid II; Brassicales; Brassicaceae; Brassica.

JOURNAL

1 (bases 1 to 804)

COMMENT

Town, C.D., Van Aken, S., Utecher, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..804

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOHRT46"

/clone_id="BOHRW74"

/note="Vector: PHOS1; Site_1: BstXI; 2-3 kb sheared

genomic DNA inserted into PHOS1 using BstXI linkers"

241 a 182 c 155 g 226 t

BASE COUNT

ORIGIN

Query Match

Best Local Similarity 2.0%; Score 32; DB 28; Length 804;

Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

835 AAGTAAACCGGTTGTATCCGATTAATC 866

DB

545 AAGTAAACCGGTTGTATCCGATTAATC 514

Search completed: January 17, 2004, 07:16:37
Job time : 3522 secs

Tue Jan 20 17:43:14 2004

us-10-022-025a-1.oligo.rng

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 13:41:32 : Search time 548 Seconds
(without alignments)
9896.297 Million cell updates/sec

Title: US-10-022-025A-1
Perfect score: 2009
Sequence: 1 aataataataataatgta.....atacagatragatctt 2009
Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 2552756 seqs, 1349719017 residues

Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

- 1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT:*
- 2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT:*
- 3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT:*
- 4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT:*
- 5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT:*
- 6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT:*
- 7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT:*
- 8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT:*
- 9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT:*
- 10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT:*
- 11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT:*
- 12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT:*
- 13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT:*
- 14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT:*
- 15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT:*
- 16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT:*
- 17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT:*
- 18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT:*
- 19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT:*
- 20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT:*
- 21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT:*
- 22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT:*
- 23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT:*
- 24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT:*
- 25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
No matches found				

Search completed: January 16, 2004, 16:01:22
Job time : 548 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 15:41:01 ; Search time 123 Seconds
(without alignments)
7209.258 Million cell updates/sec

Title: US-10-022-025A-1
Perfect score: 2009
Sequence: 1 ataataataataaagtcta.....atagagattagagatttt 2009

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 563978 seqs, 220691566 residues

Word size : 30

Total number of hits satisfying chosen parameters: 0

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

Issued Patents_NA:*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 2: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 3: /cgn2_6/prodata/1/ina/5A_COMB.seq:*
- 4: /cgn2_6/prodata/1/ina/5B_COMB.seq:*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq:*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description

No matches found

Search completed: January 16, 2004, 19:19:09
Job time : 123 secs

QY 781 TGTTCGTGAGTTCCTTAAACCGGCGTTCCTTAAACATGATGTGCTGTATTCGGACA 840
DB 781 TGTTCGTGAGTTCCTTAAACCGGCGTTCCTTAAACATGATGTGCTGTATTCGGACA 840
QY 841 AGAGTATGAGCTTGAACAAAAACATGTTGAGTTAGCTGAATGTGTGAAGAGTTATGA 900
DB 841 AGAGTATGAGCTTGAACAAAAACATGTTGAGTTAGCTGAATGTGTGAAGAGTTATGA 900
QY 901 TTTGCTCGAAGCTTGAATGAGATGATCAGCTTCCTTGGCTATCGAGTGTGATCCTCA 960
DB 901 TTTGCTCGAAGCTTGAATGAGATGATCAGCTTCCTTGGCTATCGAGTGTGATCCTCA 960
QY 961 AAGACTCCGGCTTGAATGATGATTCACACTCGTACCAAGTTAAACCGTGTGTATCCGGAT 1020
DB 961 AAGACTCCGGCTTGAATGATGATTCACACTCGTACCAAGTTAAACCGTGTGTATCCGGAT 1020
QY 1021 TATATCCGAACACCGTATCAAAACCGGATTTGCTCGTATTTGCTGACAGTTTGGCT 1080
DB 1021 TATATCCGAACACCGTATCAAAACCGGATTTGCTCGTATTTGCTGACAGTTTGGCT 1080
QY 1081 CTCCCTCCATGATGATCAATATATTCGACCCGACATATATCGCGTTCCTTGGGTATG 1140
DB 1081 CTCCCTCCATGATGATCAATATATTCGACCCGACATATATCGCGTTCCTTGGGTATG 1140
QY 1141 CACACGATTAATGATTAATTAATTTCTTAATTAATTTGTTGAATAATGCTTATGATAT 1200
DB 1141 CACACGATTAATGATTAATTAATTTCTTAATTAATTTGTTGAATAATGCTTATGATAT 1200
QY 1201 TTAGATTAAACATGAATTTGAGACTCAATGTGACGTGTGTGAATAATTAACATTT 1260
DB 1201 TTAGATTAAACATGAATTTGAGACTCAATGTGACGTGTGTGAATAATTAACATTT 1260
QY 1261 AGAAGTTTGTGTTGATCAATATTTAGTAAATTTTATGATTTATTAACAGTTTCTTAA 1320
DB 1261 AGAAGTTTGTGTTGATCAATATTTAGTAAATTTTATGATTTATTAACAGTTTCTTAA 1320
QY 1321 AAGTAAAGTACTTAAATTTTGTGATTAATTTATTAACAGATGATATTCAGAGAACG 1380
DB 1321 AAGTAAAGTACTTAAATTTTGTGATTAATTTATTAACAGATGATATTCAGAGAACG 1380
QY 1381 ACAAGTTGGGTTTATGAGTGAATCTCGTATGAGATGATTCATCCAGATATG 1440
DB 1381 ACAAGTTGGGTTTATGAGTGAATCTCGTATGAGATGATTCATCCAGATATG 1440
QY 1441 AATCAAGCTACAAACAGAGCTGATCAAGTATGCGGAAATCAAGAGCCCTGATGAAT 1500
DB 1441 AATCAAGCTACAAACAGAGCTGATCAAGTATGCGGAAATCAAGAGCCCTGATGAAT 1500
QY 1501 CTGACTTGGCTTCACTTCATATCTAACCGGCTGTGTGAAGAGATTTAGGCTTATC 1560
DB 1501 CTGACTTGGCTTCACTTCATATCTAACCGGCTGTGTGAAGAGATTTAGGCTTATC 1560
QY 1561 CTCAGAGCCCACTTCTATCATGAGCCCGTTGGCCATTAACAGACATGTTGATGATC 1620
DB 1561 CTCAGAGCCCACTTCTATCATGAGCCCGTTGGCCATTAACAGACATGTTGATGATC 1620
QY 1621 GTCTTGTTCGGGACAGGACACAGCAATGTTGAACATGTGGCCGTTATCCATGATCCAC 1680
DB 1621 GTCTTGTTCGGGACAGGACACAGCAATGTTGAACATGTGGCCGTTATCCATGATCCAC 1680
QY 1681 ACGTGTGGGTTGATCTTGTGAGTTTAAACCTGAGAGTTGCTGTGCAAAAGAGTGAAG 1740
DB 1681 ACGTGTGGGTTGATCTTGTGAGTTTAAACCTGAGAGTTGCTGTGCAAAAGAGTGAAG 1740
QY 1741 TGAAGTTTGGGTTTGGGTTGAGATTTGAGATTTGACATTTTCGGGTCTGGTCTCGGA 1800
DB 1741 TGAAGTTTGGGTTTGGGTTGAGATTTGAGATTTGACATTTTCGGGTCTGGTCTCGGA 1800
QY 1801 TTTGCCCCGGGAGAGATCTTGTGTTTACTACGTTATGTTTGAAGCGGATGATGTATC 1860
DB 1801 TTTGCCCCGGGAGAGATCTTGTGTTTACTACGTTATGTTTGAAGCGGATGATGTATC 1860
QY 1861 ATGAGTTTGAATGGGAGCGTCCAGTGTGAACGGCGTTGACTTATCTGAGAAACTGAGGC 1920

DB 1861 ATGAGTTTGAATGGGAGCGTCCAGTGTGAACGGCGTTGACTTATCTGAGAAACTGAGGC 1920
QY 1921 TTTCTTGGAGATGAGTATCTTCTTCTGCTAAATTTGCGCGGATGAGGAGTTAAATAA 1980
DB 1921 TTTCTTGGAGATGAGTATCTTCTTCTGCTAAATTTGCGCGGATGAGGAGTTAAATAA 1980
QY 1981 AAGAAGCTCATATGAGATTTAGAGATTTT 2009
DB 1981 AAGAAGCTCATATGAGATTTAGAGATTTT 2009

RESULT 2
US-09-349-385-5
; Sequence 5, Application US/09349385
; Patent No. US20020152495A1
; GENERAL INFORMATION:
; APPLICANT: Ito, Toshio
; APPLICANT: Fromm, Michael
; APPLICANT: Meyerowitz, Elliot
; TITLE OF INVENTION: PLANTS HAVING SEEDLESS FRUIT
; FILE REFERENCE: MB1-0002
; CURRENT APPLICATION NUMBER: US/09/349,385
; EARLIER FILING DATE: 1999-07-09
; EARLIER APPLICATION NUMBER: 60/115,967
; EARLIER FILING DATE: 1999-01-15
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; OTHER INFORMATION: cDNA
US-09-349-385-5

Query Match 48.6%; Score 976; DB 10; Length 1593;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 161 ATGCTAGGAAGTCTGAAGTCTTATCTTCCCTTGTGCAATGAGAGGTTCTA 220
DB 1 ATGCTAGGAAGTCTGAAGTCTTATCTTGTGCTTGTGCAATGAGAGGTTCTA 60
QY 221 AGCCAAACCACTTCCCTTCTCCCTCTCGCGTCAATCATCTGCTGCGATATCT 280
DB 61 AGCCAAACCACTTCCCTTCTCCCTCTCGCGTCAATCATCTGCTGCGATATCT 120
QY 281 CTCTCTTATGACCTATCCCGGAGACCTGAGGAGAAATGCTTGGCGGCTA 340
DB 121 CTCTCTTATGACCTATCCCGGAGACCTGAGGAGAAATGCTTGGCGGCTA 180
QY 341 ATATCCGTTATACAAACCGGAAACGTTATTCGCGTCCAAAGGCTTCCCTTGT 400
DB 181 ATATCCGTTATACAAACCGGAAACGTTATTCGCGTCCAAAGGCTTCCCTTGT 240
QY 401 GGAAGCATGCTACTATGCTCAAGCACTTACTACCGAAGATCGGTGAGCTGAG 460
DB 241 GGAAGCATGCTACTATGCTCAAGCACTTACTACCGAAGATCGGTGAGCTGAG 300
QY 461 AAATTCGAGCCAGAGGCTATGCTTTCAGCTTGAAGAGACTGCGGTATGTCAG 520
DB 301 AAATTCGAGCCAGAGGCTATGCTTTCAGCTTGAAGAGACTGCGGTATGTCAG 360
QY 521 TGCATCCCGAGCTAGCGAAAGATTTGTAATAGCCCGGTTTGTGATGACCGGCT 580
DB 361 TGCATCCCGAGCTAGCGAAAGATTTGTAATAGCCCGGTTTGTGATGACCGGCT 420
QY 581 AAGAATCGGCTTACTCATGATGTTTAAACAGAGCATTTGTTTGAACAACGGTGT 640
DB 421 AAGAATCGGCTTACTCATGATGTTTAAACAGAGCATTTGTTTGAACAACGGTGT 480
QY 641 TACTGCGAAGCGCTTCCGCTATGCTTGAACCATCTCTTATGTAACAAATCA 700

```

Db      481 TACTGCGAAGCGCTTGCCTGATCGCTTGAACCACTCTTATGACAAACAAATCGA 540
Qy      701 AGAGCGAGACCGCAACGACGATGATCTCAACCCAGATGTTGATTTCTGAAAAAG 760
Db      541 AAGCGCGAGACCGCAACGACGATGATCTCAACCCAGATGTTGATTTCTGAAAAAG 600
Qy      761 AGTAGTAACGAACCTGTTTGTGTGTGATGTTCTTAAACCGCGCTGCTTAAACAATG 820
Db      601 AGTAGTAACGAACCTGTTTGTGTGTGATGTTCTTAAACCGCGCTGCTTAAACAATG 660
Qy      821 ATGTGCTGTATATCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 880
Db      661 ATGTGCTGTATATCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy      881 ATGTGCGAAGAGATTATGATTTGCTCGGACGTTGATGATGATGATGATGATGATGAT 940
Db      721 ATGTGCGAAGAGATTATGATTTGCTCGGACGTTGATGATGATGATGATGATGATGAT 780
Qy      941 CTATCGGAGTTGATGATCTCAAGACGTCGCTGATGATGATGATGATGATGATGATGAT 1000
Db      781 CTATCGGAGTTGATGATCTCAAGACGTCGCTGATGATGATGATGATGATGATGATGAT 840
Qy      1001 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAATCAACCGGTGATTTGCTGCT 1060
Db      841 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAATCAACCGGTGATTTGCTGCT 900
Qy      1061 GATTTCGTGAGCTTTGCTCTCCCTGCAATGCTGATGATGATGATGATGATGATGATGAT 1120
Db      901 GATTTCGTGAGCTTTGCTCTCCCTGCAATGCTGATGATGATGATGATGATGATGAT 960
Qy      1121 ATCGCGTTCTTTGGG 1136
Db      961 ATCGCGTTCTTTGGG 976

RESULT 3
US-10-022-025a-2
; Sequence 2, Application US/10022025a
; Publication No. US20030092014A1
; GENERAL INFORMATION:
; APPLICANT: Coleman, John R.
; TITLE OF INVENTION: Nucleic Acid Molecules and Polypeptides for Catabolism of Abscisi
; FILE OF INVENTION: Acid
; FILE REFERENCE: 3310 0003
; CURRENT APPLICATION NUMBER: US/10/022,025A
; CURRENT FILING DATE: 2003-01-09
; PRIOR APPLICATION NUMBER: US 60/254,819
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: Patent version 3.0
; SEQ ID NO 2
; LENGTH: 1593
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1593)
US-10-022-025a-2

Query Match      48.6%; Score 976; DB 15; Length 1593;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 976; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Db      121 CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCGCGCGTTA 180
Qy      341 ATATCCGTTCAATCAAAAACCGGAAACGTTATTCGGGTCCAAAAGGCTTCCCTTTGGTT 400
Db      181 ATATCCGTTCAATCAAAAACCGGAAACGTTATTCGGGTCCAAAAGGCTTCCCTTTGGTT 240
Qy      401 GGAAGCATGTCATCATGTCAGACACTGATGCTCACGAGCATGCTGATGAGCTGAG 460
Db      241 GGAAGCATGTCATCATGTCAGACACTGATGCTCACGAGCATGCTGATGAGCTGAG 300
Qy      461 AAATTCGAGCCAAAGGCTCATGCTTCAAGCTTGAAGAGACTGCGTGAATGCTCAG 520
Db      301 AAATTCGAGCCAAAGGCTCATGCTTCAAGCTTGAAGAGACTGCGTGAATGCTCAG 360
Qy      521 TGCAATCCGACGTAAGGAAAGATTCGAATAGCCCGGTTTGTGATGACACCGGTT 580
Db      361 TGCAATCCGACGTAAGGAAAGATTCGAATAGCCCGGTTTGTGATGACACCGGTT 420
Qy      581 AAAGATCGGCTTACTCATGATGTTTAAAGAGCAATTGTTTGAACCAACGAGTGT 640
Db      421 AAAGATCGGCTTACTCATGATGTTTAAAGAGCAATTGTTTGAACCAACGAGTGT 480
Qy      641 TACTGCGAAGCGTTGCGCTGATGCTTGAACCACTCTTTAGTACAAAACAATCGA 700
Db      481 TACTGCGAAGCGTTGCGCTGATGCTTGAACCACTCTTTAGTACAAAACAATCGA 540
Qy      701 AGAGCGAGACGCAACGACGATGATCTCAAGCCAGATGATGATGATGATGATGATGATGAT 760
Db      541 AGAGCGAGACGCAACGACGATGATCTCAAGCCAGATGATGATGATGATGATGATGATGAT 600
Qy      761 AGTAGTAACGAACCTGTTTGTGCTGATGCTTAAACGGGTGCTTAAACAATG 820
Db      601 AGTAGTAACGAACCTGTTTGTGCTGATGCTTAAACGGGTGCTTAAACAATG 660
Qy      821 ATGTGCTGTATATCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 880
Db      661 ATGTGCTGTATATCGGACAGAGATGATGATGATGATGATGATGATGATGATGATGAT 720
Qy      881 ATGTGCGAAGAGATTATGATTTGCTCGGACGTTGATGATGATGATGATGATGATGAT 940
Db      721 ATGTGCGAAGAGATTATGATTTGCTCGGACGTTGATGATGATGATGATGATGATGAT 780
Qy      941 CTATCGGAGTTGATGATCTCAAGACGTCGCTGATGATGATGATGATGATGATGATGAT 1000
Db      781 CTATCGGAGTTGATGATCTCAAGACGTCGCTGATGATGATGATGATGATGATGATGAT 840
Qy      1001 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAATCAACCGGTGATTTGCTGCT 1060
Db      841 AACCGGTTTGTATCCCGGATTAATCCGAAACCGTAATCAACCGGTGATTTGCTGCT 900
Qy      1061 GATTTCGTGAGCTTTGCTCTCCCTGCAATGCTGATGATGATGATGATGATGATGATGAT 1120
Db      901 GATTTCGTGAGCTTTGCTCTCCCTGCAATGCTGATGATGATGATGATGATGATGAT 960
Qy      1121 ATCGCGTTCTTTGGG 1136
Db      961 ATCGCGTTCTTTGGG 976

```

Search completed: January 16, 2004, 21:27:40
Job time : 663 secs

Qy 221 AGCAAAACCAACCTTGGCTTCTCCCTGCTGCGCTGCAATCATGCTGCGCATATCT 280
Db 61 AGCAAAACCAACCTTGGCTTCTCCCTGCTGCGCTGCAATCATGCTGCGCATATCT 120
Qy 281 CTCTTTTATGACCTATCCCGGTGACCTGCTTGGGGGAAATACCTCTTCGCGCGTTA 340

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

CM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 00:13:09 ; Search time 450 Seconds
(without alignments)
9536.011 Million cell updates/sec

Title: US-10-022-025a-2

Perfect score: 1593

Sequence: 1 atgctacgaactcgaag.....tcgcgcgtagcgcagcttaa 1593

Scoring table:

IDENTITY NUC

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

N_Geneseq_19Jun03:*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*
25: /SIDSI/gcgdata/geneseq/geneseq-emb1/NA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	976	61.3	1902	25	ABX10941
2	975.6	61.2	2005	21	AACT7902
3	794	49.8	7869	25	ABX10942
4	651.6	40.9	1611	20	AAK60776
5	599.6	37.6	1764	25	ABT23239
6	585.8	36.8	1905	25	ABT23241
7	355	22.3	1799	20	AAK55600
8	352	22.1	1930	25	ABT23246

9	316.6	19.9	1934	25	ABT23240	Seed development e
10	290.6	18.2	1687	25	ABT23249	Seed development e
11	286	18.0	1545	25	ABT23284	Seed development e
12	283	17.8	1758	25	ABT23283	Seed development e
13	278.2	17.5	1597	25	ABT23237	Seed development e
14	275.2	17.3	1585	25	ABT23282	Seed development e
15	274.4	17.2	1539	25	ABT23238	Seed development e
16	265.8	16.7	2924	25	ABT23242	Seed development e
17	234	14.7	1146	25	ABT23235	Seed development e
18	233.4	14.7	1578	25	ABT23228	Seed development e
19	213.4	13.4	1131	25	ABT23233	Seed development e
20	194.8	12.2	1557	25	ABT23285	Seed development e
21	189.2	11.9	884	25	ABT23247	Seed development e
22	156.4	9.8	442	25	ABT23248	Seed development e
23	155.6	9.8	17201	25	ABT23229	Seed development e
24	144.2	9.1	1704	20	AAK60781	Soybean cytochrome
25	141.8	8.9	1789	18	AAK94655	Pectunia flavonoid
26	135.6	8.5	1667	18	AAK94664	Gentiana flavonoid
27	129.6	8.1	1708	20	AAK81873	DNA encoding a P45
28	129.6	8.1	1708	21	AAA88681	Sweetgum cytochrom
29	129.6	8.1	1708	22	AAH1125	L. styraciflua ang
30	126	7.9	778	25	ABT23236	Seed development e
31	124.4	7.8	1660	18	AAK94661	Chrysanthemum flav
32	124.4	7.8	1748	18	AAK94660	Rose flavonoid 3'-
33	121.2	7.6	1745	18	AAK94656	Carnation flavonol
34	119.4	7.5	1527	24	ABK12518	DNA encoding p-con
35	119.4	7.5	1549	24	ABK12520	DNA construct for
36	117.8	7.4	1527	24	ABK12519	DNA encoding modif
37	117.4	7.4	1711	18	AAK94657	Snapdragon flavono
38	113.8	7.1	1214	18	AAK94655	Isianthus flavono
39	111.8	7.0	1121	21	AAK36642	Arabidopsis thalia
40	111.8	7.0	1815	18	AAK94662	Toronia flavonoid
41	109.2	6.9	1824	18	AAK94663	Morning glory flav
42	106.8	6.7	1770	22	AAK30281	Perilla flavone sy
43	106.8	6.7	1933	25	ABX93222	cDNA encoding cot
44	105.6	6.6	457	25	ABT23244	Seed development e
45	104	6.5	618	22	AAK87696	Peppermint plant o

ALIGNMENTS

RESULT 1	ABX10941	Standard; cDNA, 1902 BP.
ID	ABX10941	Standard; cDNA, 1902 BP.
XX	ABX10941;	
AC	ABX10941;	
XX	16-APR-2003	(first entry)
DT	16-APR-2003	(first entry)
XX	CDNA encoding Arabidopsis thaliana cytochrome P450 protein.	
DE	Arabidopsis; cytochrome P450; CP450; plant; parthenocarpic; fruit;	
XX	enlarged fruit; transgenic; vegetable size; grain size; gene; ss;	
KW	leaf size; flower size; agricultural yield; male-sterile plant.	
KW		
XX	Arabidopsis thaliana.	
OS	Arabidopsis thaliana.	
XX		
PH	Key	Location/Qualifiers
FT	CDS	69..1673
FT		/tag= a
FT		/product= "A. thaliana cytochrome P450"
XX	US2002152495-A1.	
PN	17-OCT-2002.	
XX	09-JUL-1999;	99US-0349385.
PF	09-JUL-1999;	99US-0349385.
XX	15-JAN-1999;	99US-115967P.
PR	15-JAN-1999;	99US-115967P.
XX	(ITON/) ITO T.	
XX	(FROM/) FROM M.	
PA		

PA (MEYE/) MEYEROWITZ E.
 XX Ito T, Fromm M, Meyerowitz E;
 XX WPI, 2003-198390/19.
 DR P-PSDB; AB575612.
 XX
 XX New polynucleotides encoding a cytochrome P450 polypeptide, useful for
 PT producing plants with enlarged or parthenocarpic fruits, in increasing
 PT plant tissue size to increase agricultural yields, or producing
 PT male-sterile plants
 XX
 XX Claim 4, Page 12-13, 31pp, English.
 PS
 XX This invention relates to an isolated polynucleotide encoding a
 CC cytochrome P450 polypeptide which when expressed in a plant produces at
 CC least one phenotype selected from parthenocarpic fruit and enlarged
 CC fruit. The invention also discloses a recombinant construct comprising
 CC the novel polynucleotide which when expressed in a plant produces a
 CC plant with a parthenocarpic fruit or enlarged fruit. Also disclosed is
 CC a transgenic plant which expresses a cytochrome P450 or comprising
 CC an isolated polynucleotide encoding a cytochrome P450 polypeptide which
 CC when expressed in a plant produces a plant with a parthenocarpic fruit
 CC or enlarged fruit compared with a plant lacking the isolated
 CC polynucleotide. The cytochrome P450 polypeptide and polynucleotide of
 CC the invention is useful in the production of plants with seedless,
 CC enlarged, or parthenocarpic fruits, including vegetable or grain size,
 CC leaf size or flower size. The sequences may also be used in increasing
 CC plant tissue size to increase agricultural yields of plants, to produce
 CC male-sterile plants and to screen for compounds that control
 CC parthenocarpic or fruit size in plants. The present sequence represents
 CC the cDNA sequence encoding the Arabidopsis thaliana cytochrome P450
 CC protein of the invention.
 XX
 XX Sequence 1902 BP, 507 A; 448 C; 439 G; 508 T; 0 other;
 SQ
 Query Match 61.3%; Score 976; DB 25; Length 1902;
 Best Local Similarity 78.0%; Pred. No. 0;
 Matches 128; Conservative 0; Mismatches 320; Indels 30; Gaps 4;

DB 555 AAATGCGCTTAATCATCTTTTCAGCCGAAACAGATTAAACGTTCCGAAACGACGAGA 614
 QY 559 CGAGTATCTCAAGCAGATGGTGAATTTCTTGAAGAAACAGATAGTAACGA---CC 615
 DB 615 AGGTGATGCGAATCAAAATCGTGAAGTGTCTCAAAAACAGATACCAAGGCTGC 674
 QY 616 TGTGTTGTCGTGAGTGGCTTAAACGGCGTGCCTTAAACAATGATGCTCTGATTC 675
 DB 675 TGTTCGCAAGTGAAGCTTGAACAAACGGCATGCTTAATACATATGCTGCTGTTTC 734
 QY 676 GGAAGAGATGAGCTGAAAAAACATG-----TTAGCTAGTGAATGCTC 726
 DB 735 GAAAGAAATTCAGAGCTTGAAGAAAGATGAAGAGTGAAGTGAAGTGAATGGTG 794
 QY 727 GAAGAAAGTTATGATTTGCTCGAAGCTTGAATTTGAGATGATCACTTCTTGGCTATCG 786
 DB 795 GAAGAAAGTTATGATTTTACTCGGTACAGTAATTTGACCGATCATCTCCATGCTCTCT 854
 QY 787 GACTTGAATCCCAAGACTCGGTCTAGATGTTCCACCTGTGCAAAAGGTAACCGG 846
 DB 855 GAATTTGATCTCAAGAAATCGGTCTAATGCTTAACTGCTCCAAAAGTAACCGG 914
 QY 847 TTTGTATCCCGATTTATTCGAAACACCGTAATCAACCGGTGATTTGCTGTGATTC 906
 DB 915 TTTGTGAACCGGATTTATCTCTGACACCGGTAACAACTGTGACTCACCGAGTACTTC 974
 QY 907 GTGAGCTTTTGTCTCCCTCCATGGTTCAATTAATTTCCGACCCGGAACATTAATCGCC 966
 DB 975 GTTGAAGTATGCTCTCTCGATGCTCTGATTAATTTCCGACCTTAATTCATTCGCC 1034
 QY 967 GTCTTTGGAGATGATTAATTCAGAGAGACACAGTTCGGCTTAAATCGATGATC 1026
 DB 1035 GTTCTATGGAAATGATTAATTCAGAGAGACACAGTTCGGCTTAAATCGATGATC 1094
 QY 1027 CTGGCTATGATGCTCTTATTCAGATATTCATCAACGGTATCAAAACGATCGATCA 1086
 DB 1095 CTGGCTATGATGCTCTTATTCAGATATTCATCAACGGTATCAAAACGATCGATCA 1154
 QY 1087 GTAGTGGGAAATCAAGAGCTTATGATGATGATGATGATGATGATGATGATGATG 1146
 DB 1155 ATGCTGGAAGATCAAGAGCTGTCGAAGATGATGATGATGATGATGATGATGATG 1214
 QY 1147 GCTGTGTGAAGAAAGTATGAGCTTATCTCTCAAGGCCATTTATCATATGAGCCGT 1206
 DB 1215 GCTGTGTGAAGAAAGTATGAGCTTATCTCTCAAGGCCATTTATCATATGAGCCGT 1274
 QY 1207 TTGGCCATTAACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1266
 DB 1275 TTAGCAATCAACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATG 1334
 QY 1267 GTGAACATGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1326
 DB 1335 GTGAACATGTGGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATG 1394
 QY 1327 CTTGAAGCTGCTGGGAAAGAAAGGATGATGATGATGATGATGATGATGATGATGATG 1386
 DB 1395 CCGGAAGCTTGTAGCAAGAAAGGATGATGATGATGATGATGATGATGATGATGATG 1454
 QY 1387 AGACTTCACCTTTCGGGTGGGTGCTGCGATTTGCCCCGGGAAAGAACTTGGTTTACT 1446
 DB 1455 AGCTTGCACCGTTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGGTCCGG 1514
 QY 1447 ACCGTTATGTTTGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1503
 DB 1515 ACCGTTATGTTTGAAGGAGATGATGATGATGATGATGATGATGATGATGATGATG 1574
 QY 1504 GTTAACGCGCTTGAATTTATCTGAAGAACTGAGGCTTTCTTTCGAGATGAGCTAAATCTCT 1563
 DB 1575 GAGAAGACCGTTGATCTTGTCCGAAAGACTGAGGCTCTGATGATGATGATGATGATGAT 1634
 QY 1564 CTTGCTAAATTTGCGCCGTAGCGGCAATT 1591

```
DB      1635 GCTGCTAATTACGCCCGACGCGCACT 1662

RESULT 2
AAC47902
ID      AAC47902 standard; DNA; 2005 BP.
XX
XX      AAC47902;
AC
XX      18-OCT-2000 (first entry)
DT
XX
XX      Arabidopsis thaliana DNA fragment SEQ ID NO: 55538.
DE
XX
XX      Hybridisation assay; genetic mapping; gene expression control;
XX      protein identification; signal transduction pathway;
XX      metabolic pathway; promoter; termination sequence; ss.
OS
XX      Arabidopsis thaliana.
XX      EP1033405-A2.
PN
XX
XX      06-SEP-2000.
PD
XX
XX      25-FEB-2000; 2000EP-0301439.
PF
XX
XX      25-FEB-1999; 99US-0121825.
XX      05-MAR-1999; 99US-0123180.
PR      09-MAR-1999; 99US-0123548.
PR      23-MAR-1999; 99US-0125788.
PR      25-MAR-1999; 99US-0126264.
PR      29-MAR-1999; 99US-0126785.
PR      01-APR-1999; 99US-0127462.
PR      06-APR-1999; 99US-0128224.
PR      08-APR-1999; 99US-0128714.
PR      16-APR-1999; 99US-0129845.
PR      19-APR-1999; 99US-0130077.
PR      21-APR-1999; 99US-0130419.
PR      23-APR-1999; 99US-0130510.
PR      28-APR-1999; 99US-0130891.
PR      30-APR-1999; 99US-0132048.
PR      30-APR-1999; 99US-0132407.
PR      04-MAY-1999; 99US-0132484.
PR      05-MAY-1999; 99US-0132485.
PR      06-MAY-1999; 99US-0132486.
PR      06-MAY-1999; 99US-0132487.
PR      07-MAY-1999; 99US-0132863.
PR      11-MAY-1999; 99US-0134256.
PR      14-MAY-1999; 99US-0134218.
PR      14-MAY-1999; 99US-0134219.
PR      14-MAY-1999; 99US-014221.
PR      14-MAY-1999; 99US-014370.
PR      18-MAY-1999; 99US-0134768.
PR      19-MAY-1999; 99US-014941.
PR      20-MAY-1999; 99US-0135124.
PR      21-MAY-1999; 99US-0135353.
PR      24-MAY-1999; 99US-0135629.
PR      25-MAY-1999; 99US-0136021.
PR      27-MAY-1999; 99US-0136382.
PR      28-MAY-1999; 99US-0136782.
PR      01-JUN-1999; 99US-0137222.
PR      03-JUN-1999; 99US-0137502.
PR      04-JUN-1999; 99US-0137724.
PR      07-JUN-1999; 99US-0138094.
PR      08-JUN-1999; 99US-0138549.
PR      10-JUN-1999; 99US-0138847.
PR      14-JUN-1999; 99US-0139119.
PR      16-JUN-1999; 99US-0139452.
PR      16-JUN-1999; 99US-0139453.
PR      17-JUN-1999; 99US-0139492.
PR      18-JUN-1999; 99US-0139454.
PR      18-JUN-1999; 99US-0139455.

PR      18-JUN-1999; 99US-0139456.
PR      18-JUN-1999; 99US-0139457.
PR      18-JUN-1999; 99US-0139458.
PR      18-JUN-1999; 99US-0139459.
PR      18-JUN-1999; 99US-0139460.
PR      18-JUN-1999; 99US-0139461.
PR      18-JUN-1999; 99US-0139462.
PR      18-JUN-1999; 99US-0139463.
PR      18-JUN-1999; 99US-0139750.
PR      18-JUN-1999; 99US-0139753.
PR      21-JUN-1999; 99US-0139817.
PR      22-JUN-1999; 99US-0139899.
PR      23-JUN-1999; 99US-0140353.
PR      23-JUN-1999; 99US-0140354.
PR      24-JUN-1999; 99US-0140695.
PR      28-JUN-1999; 99US-0140823.
PR      29-JUN-1999; 99US-0140981.
PR      30-JUN-1999; 99US-0141287.
PR      01-JUL-1999; 99US-0141842.
PR      01-JUL-1999; 99US-0142154.
PR      02-JUL-1999; 99US-0142055.
PR      06-JUL-1999; 99US-0142390.
PR      08-JUL-1999; 99US-0142803.
PR      09-JUL-1999; 99US-0142920.
PR      12-JUL-1999; 99US-0142977.
PR      13-JUL-1999; 99US-0143542.
PR      14-JUL-1999; 99US-0143624.
PR      15-JUL-1999; 99US-0144005.
PR      16-JUL-1999; 99US-0144085.
PR      16-JUL-1999; 99US-0144086.
PR      19-JUL-1999; 99US-0144325.
PR      20-JUL-1999; 99US-0144352.
PR      20-JUL-1999; 99US-0144632.
PR      20-JUL-1999; 99US-0144864.
PR      21-JUL-1999; 99US-0144864.
PR      21-JUL-1999; 99US-0145086.
PR      21-JUL-1999; 99US-0145087.
PR      22-JUL-1999; 99US-0145088.
PR      22-JUL-1999; 99US-0145089.
PR      22-JUL-1999; 99US-0145192.
PR      22-JUL-1999; 99US-0145193.
PR      23-JUL-1999; 99US-0145115.
PR      23-JUL-1999; 99US-0145218.
PR      26-JUL-1999; 99US-0145224.
PR      26-JUL-1999; 99US-0145276.
PR      27-JUL-1999; 99US-0145913.
PR      27-JUL-1999; 99US-0145918.
PR      27-JUL-1999; 99US-0145919.
PR      28-JUL-1999; 99US-0145951.
PR      02-AUG-1999; 99US-0146386.
PR      02-AUG-1999; 99US-0146388.
PR      02-AUG-1999; 99US-0146389.
PR      03-AUG-1999; 99US-0147038.
PR      04-AUG-1999; 99US-0147204.
PR      04-AUG-1999; 99US-0147302.
PR      05-AUG-1999; 99US-0147192.
PR      05-AUG-1999; 99US-0147260.
PR      06-AUG-1999; 99US-0147303.
PR      06-AUG-1999; 99US-0147416.
PR      09-AUG-1999; 99US-0147493.
PR      09-AUG-1999; 99US-0147935.
PR      10-AUG-1999; 99US-0148171.
PR      11-AUG-1999; 99US-0148319.
PR      12-AUG-1999; 99US-0148341.
PR      13-AUG-1999; 99US-0148665.
PR      13-AUG-1999; 99US-0148684.
PR      16-AUG-1999; 99US-0149368.
PR      17-AUG-1999; 99US-0149175.
```


PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149920.
PR 23-AUG-1999; 99US-0149920.
PR 25-AUG-1999; 99US-0150586.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152353.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159283.
PR 13-OCT-1999; 99US-0159284.
PR 13-OCT-1999; 99US-0159285.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160747.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 61.2%; Score 975.6; DB 21; Length 2005;
Best Local Similarity 77.9%; Pred. No. 0;
Matches 1237; Conservative 1; Mismatches 320; Indels 30; Gaps 4;

QY 19 AGCTCCCTATCTTGGCCCTTTCCTCAATGAGGTTCTTCAAGCAACCACTTGGC 78
DB 199 AGTAGCTACTTTGGCCCTTTCCTCAATGAGGTTCTTCAAGCAACCACTTGGC 258
QY 79 TTCTCCCTCTCCGCGGTACCAATCATCTGGCTGCCATATCTCTTTTATGACCTAT 138
DB 259 CTCTCTCTCTCGAGCTCCCTAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCAT 318

DB 319 CCGGAGAGCCGCACTGGGAAAAGTACTTCTCCACCGC-----CGCGGT 363
QY 199 ACCGGAACGTATTCCCGGTCCAAAGGCTTCCCTTTGGTGAAGATGTACTATG 256
DB 364 CAACCACTGTAACCGGAGCCAGAGGCTTACCTTTTGTGGAAGATGTCTTATG 423
QY 259 TCAAGCACTTAGCTACCGAGATGCTGATGAGTGAATTCGAGCCAGAGG 338
DB 424 TCAACACTGCTGCTACCGTGTGACCGCCCAACCGGAGAAATTAGAGCCGGA 483
QY 319 CTGATGCTTTCAGCTTGAAGAGATGCGGAGTTCGACGTCATTCGCGATGCG 378
DB 484 TTAATGCGTTTATGTTTGGAGAACTGCGTATTCGACGTCATTCGATGTAGT 543
QY 379 AAAGATTTGATATAGCCCGGTTTGTGATGACCGGTTAAGATGCGCTTACTCA 438
DB 544 AAAGATTTCTAAGACGTCGCGTTTGTGATGACCGGTTAAGATGCGCTTACTCC 603
QY 439 CTGATGTTTACAGCAATGCTTTTGCACACAGGTTTTCGAGCGAGCTTCGCG 498
DB 604 CTGATGTTTACAGCAATGCTTTTGCCTCTTACGCGCTTTTACGCGAGCTTGA 663
QY 499 CGTATGCTTGAACCAATCTCTTATGTAACAACAATCAGAGAGCGGAGCGA 558
DB 664 AAATGCGCTCTAATCATCTTTTCAGCCGAAACAGATTAAAGCTTCGAAAG 723
QY 559 CGAGTATCTCAAGCAAGATGCTTGAATTTCTTGAACAGAGTATTAAGAA---CCG 615
DB 724 AGCTGATGCGCAATCAATGTAAGTGTCTCAACAGAGTATCAACAAAGCTTC 783
QY 616 TGTGTTGTTGTTGAGTGTCTTAAACGCGGTGCTTAAACATGATGTCTGTATTC 675
DB 784 TGTGTTGCGACGTTGATGATCAACAGCGGCTTATTAACAGATGTCTGTATTC 843
QY 676 GGACAAAGTATGAGCTTGAACAAACCATG-----TTGATGATGATGAAATGTC 726
DB 844 GGAAGAGTATGAGCTTGAACAGAGAGATGAAGATGAAGATGATGATGATGTC 903
QY 727 GAAAGAGTATGATTTGCTGGAACGTTGATGATGATGATGATGATGATGATG 786
DB 904 GAAAGAGTATGATTTGCTGGAACGTTGATGATGATGATGATGATGATGATG 963
QY 787 GAGTTGATCTCAAGAGCTCGGCTGATGATGATGATGATGATGATGATGATG 846
DB 964 GATTTGATCTCAAGAGTCCGCTGATGATGATGATGATGATGATGATGATG 1023
QY 847 TTTGATCTCCGATTTATTCGGAACGTTGATGATGATGATGATGATGATG 906
DB 1024 TTTGATCTCCGATTTATTCGGAACGTTGATGATGATGATGATGATGATG 1083
QY 907 GTGACGTTTGTCTCCCTGATGATGATGATGATGATGATGATGATGATG 966
DB 1084 GTTGAAGTATGCTCTCTGATGATGATGATGATGATGATGATGATGATG 1143
QY 967 GTTCTTTGGAGATGATTTGAGAGAACAGACATGTCGCTTATTCGATGATG 1026
DB 1144 GTTCTTTGGAGATGATTTGAGAGAACAGACATGTCGCTTATTCGATGATG 1203
QY 1027 CTGCTAGGATGATGCTTATTCAGATGATGATGATGATGATGATGATGATG 1086
DB 1204 CTGCTAGGATGATGCTTATTCAGATGATGATGATGATGATGATGATGATG 1263
QY 1087 GTATGCGGAATCAAGAGCTGATGATGATGATGATGATGATGATGATG 1146
DB 1264 ATGATGCGGAATCAAGAGCTGATGATGATGATGATGATGATGATGATG 1323
QY 1147 GCTATGATGAAAGATGATGATGATGATGATGATGATGATGATGATG 1306
DB 1324 GCTATGATGAAAGATGATGATGATGATGATGATGATGATGATGATG 1383
QY 1207 TTGCGATTAACAGACATGATGATGATGATGATGATGATGATGATGATG 1266

Db 1384 TTAGCAATCACAGACAGATCATGACGGTCGTCGTTCCGCGGAGCAACGCGCATG 1443
 Qy 1267 GTGAACATCTGGGCGCGTATCGCATATCCACACGCTGGGTTGATCTTTGGAGTTTAA 1326
 Db 1444 GTGAACATCTGGGCGCGTATCGCATATCCACACGCTGGGAGATTCGTTGGAGTTTAA 1503
 Qy 1327 CCTGAGAGGTCGTGGCAAAAGAGTGAAGTGAATTCGTTCTTGGGTCGATTTG 1386
 Db 1504 CCGAAGCGTTTGTAGCCAAAGAGTGAAGTGAATTCGTTCTTGGGTCGATTTG 1563
 Qy 1387 AGACTTGCACCTTTGGGTCGCGGTCTGTCGATTTTCCCGGGAAGAACTTGGTTTACT 1446
 Db 1564 AGGCTTGCACCGTTCGGGTCGCGGTCTGTCGATTTTCCCGGGAAGAACTTGGTTTAC 1623
 Qy 1447 ACCGTATGTTTGAACGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1503
 Db 1624 ACCGTATGTTTGAACGCGCATGATGATGATGATGATGATGATGATGATGATGATGAT 1683
 Qy 1504 GGTACGCGGTCGATTTATCTGAAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1563
 Db 1684 GAGAAAGACCGTTCGATCTGTCGGAAGAACTGAGGCTCTCGTGTGAGATGGCTAATCCT 1743
 Qy 1564 CCGCTAAATTCGCGCGTGAAGCGCAGTT 1591
 Db 1744 GCTCTAAATTCGCGCGCAGCAGTT 1771
 RESULT 3
 ABX10942
 ID ABX10942 standard; DNA; 7869 BP.
 XX ABX10942;
 AC
 XX
 XX 16-APR-2003 (first entry)
 DE Arabidopsis thaliana cytochrome P450 gene sequence.
 XX
 XX Arabidopsis; cytochrome P450; cp450; plant; parthenocarpic; fruit;
 KW enlarged fruit; transgenic; vegetable size; grain size; gene; ds;
 KW leaf size; flower size; agricultural yield; male-sterile plant.
 XX
 OS Arabidopsis thaliana.
 XX
 PN US2002152495-A1.
 PD 17-OCT-2002.
 XX
 XX 09-JUL-1999; 99US-0349385.
 PF
 XX 15-JAN-1999; 99US-115967P.
 PR
 XX (ITOT/) ITO T.
 PA (FROM/) FROM M.
 PA (MEYE/) MEYEROWITZ E.
 XX
 XX Ito T, Fromm M, Meyerowitz E;
 PI
 XX
 DR WPI; 2003-198390/19.
 XX
 PT New polynucleotides encoding a cytochrome P450 polypeptide, useful for
 PT producing plants with enlarged or parthenocarpic fruits, in increasing
 PT plant tissue size to increase agricultural yield, or producing
 PT male-sterile plants -
 XX
 PS Claim 4; Page 13-16; 31pp; English.
 XX
 CC This invention relates to an isolated polynucleotide encoding a
 CC cytochrome P450 polypeptide which when expressed in a plant produces at
 CC least one phenotype selected from parthenocarpic fruit and enlarged
 CC fruit. The invention also discloses a recombinant construct comprising
 CC the novel polynucleotide which when expressed in a plant produces a
 CC plant with a parthenocarpic fruit or enlarged fruit. Also disclosed is
 CC a transgenic plant which expresses a cytochrome P450 or comprising

CC an isolated polynucleotide encoding a cytochrome P450 polypeptide which
 CC when expressed in a plant produces a plant with a parthenocarpic fruit
 CC or enlarged fruit compared with a plant lacking the isolated
 CC polynucleotide. The cytochrome P450 polypeptide and polynucleotide of
 CC the invention is useful in the production of plants with seedless,
 CC enlarged, or parthenocarpic fruits, including vegetable or grain size,
 CC leaf size or flower size. The sequences may also be used in increasing
 CC plant tissue size to increase agricultural yield of plants, to produce
 CC male-sterile plants and to screen for compounds that control
 CC parthenocarpic or fruit size in plants. The present sequence represents
 CC the Arabidopsis thaliana cytochrome P450 genomic DNA sequence
 CC of the invention.
 XX
 SQ Sequence 7869 BP; 2628 A; 1415 C; 1374 G; 2452 T; 0 other;
 Query Match 49.8%; Score 794; DB 25; Length 7869;
 Best Local Similarity 70.3%; Pred. No. 9,5e-259; Indels 202; Gaps 5;
 Matches 1238; Conservative 0; Mismatches 320;
 Qy 19 AGCTCTTAATCTTGGCTTTTGTCCAAATGACGAGTCTAAGCCAAACCTTGGC 78
 Db 2154 AGTACCTTACTTTGGCTCTTGTCCAAATGATGATGATGATGATGATGATGATGATGAT 2213
 Qy 79 TTTCTCCTCTCGCCGTACAGATCATGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCGTCG 138
 Db 2214 CTCTCTCTCTCGTACGCTCCCTAGCTCTCTCGCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2273
 Qy 139 CCGGTGACCTGCTGGGGGAAATACCTCTGCGCGCGGTTAATTCGGTTCATCAAA 198
 Db 2274 CCGGAGACCGCATGGGAAATACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2318
 Qy 199 ACCGAAACGTTATTCGCGTCGCAAAAGGCTTCCCTTGTGGAACATGTCATG 258
 Db 2319 CAACACCGTATACCGGCGCAAGAGCTTACCTTTGCGGACGATGTCATG 2378
 Qy 259 TCAGCACTTACGTCACCGACGATGCGTATGCGTATGCGTATGCGTATGCGTATGCGTATGCG 318
 Db 2379 TCAGCACTTACGTCACCGACGATGCGTATGCGTATGCGTATGCGTATGCGTATGCGTATGCG 2438
 Qy 319 CTATGCTTACGTTAGAGAGTCTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCT 378
 Db 2439 TTAAGGCTTATGTTTGGGAAACTCGGTCGATGCTGATGCTGATGCTGATGCTGATGCT 2498
 Qy 379 AAAGGATTCGATAGCCGCTTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 438
 Db 2499 AAAGGATTCGATAGCCGCTTTTCTGATGATGATGATGATGATGATGATGATGATGATGAT 2558
 Qy 439 CTGATGTTTAAACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 498
 Db 2559 CTGATGTTTAAACAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2618
 Qy 499 CGTATGCTTGAACCATCTCTTTTGTACAAACAAATCAGAAAGCCGGAACGCAACGA 558
 Db 2619 AAATCGGCTTAAACCATCTCTTTTGTACAAACAAATCAGAAAGCCGGAACGCAACGA 2678
 Qy 559 CGAGTATCTCAAGCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 615
 Db 2679 AGCGTATGCGCATATCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2738
 Qy 616 TGTGTT 675
 Db 2739 TGTGTT 2798
 Qy 676 GGAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 726
 Db 2799 GGAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2858
 Qy 727 GAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 786
 Db 2859 GAAGAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2918
 Qy 787 GAGTTGATCTCAAGATCTCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 846

Db 313 AGAGGCTACCCCTTTCATTTGGAGATGAGTCTCATGATCC---CTGCGACACACCGCT 369
 Qy 283 ATGCTGATGACAGCTGAGAAATTCGAGCCAGAGGCTCATGCTTTTCAGCTTAGAGAG 342
 Db 370 ATTGCTGCGGCGGGGAGAGATGCAAGCCACAGGCTCATGGCTTTTTCATGGGTGAC 429
 Qy 343 ACTGCGGTGATCGTCACTGCAATCCGACGTAGCGAAAGATTTGTAATAGCCGGGT 402
 Db 430 ACAAGCGCCATAGTAAGTGAACCCGATGCTGCTTAAGATTTCTCATATGTTCCACT 489
 Qy 403 TTGTCGATGACCGGTTTAAAGATGGGCTTCACTGATGTTTAAAGAGCAATGGT 462
 Db 490 TTGCTATGCTCCCATTAAGATGATGCTTAAGCTCATGTTCAACCGCGCATCGG 549
 Qy 463 TTGCGACACACGCGTGTCTTACGCGCAACGCTTGCCTGATCGCTTGAACCATCTCTT 522
 Db 550 TTGCGCCCTTACGCGCTCTACTGCGCTTACCTCGCGCATCGCGCATCGCATCTTTC 609
 Qy 523 AGTACAAACAAATCAGAGAGCGGACGCAAGCAAGATGATCTCAAGCCAGATGTT 582
 Db 610 TGCCCAACCAATCAAGCTCCGAGCTCAGCGCTTAATTCGCCCAATGACA 669
 Qy 583 GAGTTCTTGAAGAAAGAGATGATGACAAACCTGTTTGTGCGAGTGTCTTAAACG 642
 Db 670 AACTCATTCGAAATCAACGTTGACGCGGCTTTCGGAATCGCAGCGTGTCAAGAGA 729
 Qy 643 GCGTGTCTTAAACAAGATGATGCTCTGATTCGGAACAAGATGATGCTTGAAGAAAC 702
 Db 730 GCGTCACTGACCAACATGATGATGCTGCTGTTGGAACAAAGTCAACCTTGAAGATA 789
 Qy 703 CA-----TGTGAGTACGTGAATGCTGCAAGAGGTTATGATTTGCTCGAACG 753
 Db 790 AACACCGAATGACAGAGCTATCCATGTTGTGAACAAGGCTATGATGTTGGGAC 849
 Qy 754 TTGAATGACTGATCACTTCTGCTTGGCTATCGAGTTTGTCTTCAAAAGCTCGGTCT 813
 Db 850 CTTAATGCGGAGACCAATCCCTTCTGTAAGACCTTGAACCTTGAACGAAATCCGGTTC 909
 Qy 814 AGATGTTCCACACCTGATCAACAAAGTAAACGGTTTGTATCCCGATTTATCCGAAC 873
 Db 910 ACGTCTCAATTAATGCTCCCTCAAGTAAACGGTTCGTGTTCAATCAATCGGACAC 969
 Qy 874 CGTATCAAAACCGGTGATTTGCTGCTGATTTCTGCAAGTTTGTCTCTCTCCATGCT 933
 Db 970 CAGGCGGACACCAACCAACCAACCGCATTTGCTCATGTTTGTCTCTCTCTCCAAAGT 1029
 Qy 934 TCAGATTAATATCCGACCGGACCAATATGCGCGTCTTTGGAGATGATATTCAGAGA 993
 Db 1030 CCGGTAATATGCTCACTCCGACATGATGCTGCTCTCTGGAATATATTTTAAAGGGG 1089
 Qy 994 ACAAGACAGATGCGGTCTTATGATGATGATCTCTGCTAGATGCTCTTCAATCAAT 1053
 Db 1090 ACCGACAGGTGCGGTTTGTGATGATGATGATGATGATGATGATGATGATGATGAT 1149
 Qy 1054 ATGCAATCAACGCTCAAAAGAGGCTGATCAAGATGCGGAAATCAAGAGCCCTAGAT 1113
 Db 1150 GTGCAAAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1206
 Qy 1114 GATCTGATCTTGGCTTCACTTCCATATCTAAACGCTGATGATGATGATGATGATGAT 1173
 Db 1207 GAGGAGGTGCTGCGGACGCGCTATCTTGGCGGCTGATGATGATGATGATGATGAT 1266
 Qy 1174 CATCTCCAGGCTCTTCTATCATGATGATGATGATGATGATGATGATGATGATGAT 1233
 Db 1267 CACCGCGCGGCTCTCTCTGATGATGATGATGATGATGATGATGATGATGATGAT 1326
 Qy 1234 GGTGCTTCTTCCGACGAGGACCAAGCAATGATGATGATGATGATGATGATGATGAT 1293
 Db 1327 GGGTATCAAGTCTGCGGAGCAACCGTATGATGATGATGATGATGATGATGATGAT 1386
 Qy 1294 CCAACAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1353
 Db 1387 CCGAGGTGTGCTGACCACTTGAATCAAGCCGAGAGGTGATGATGATGATGATGAT 1443

Qy 1354 GAGTGAAGTTTTCGTTCTTGGGTCCGATTTTGAAGACTTGCACCTTCCGGTCCGCT 1413
 Db 1444 --AACGAGTTTCTGTTTCCGGTCCGATCTGAGATCGCTCCATTCGGTCCGCT 1500
 Qy 1414 CGGATTTCCCGGAGAAATCTTGTATTTTACCTGATTTGTTTGAAGCGGATGATG 1473
 Db 1501 AGAATATCCCGGAGAGACTTGGTTGAGACCGTAAACCTTCTGGGTGGCTT 1560
 Qy 1474 TTACATGATTTGAATGGAGCGCTCCGATGATGATGATGATGATGATGATGATGAT 1533
 Db 1561 TTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1620
 Qy 1534 AGGCTTCTGCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1593
 Db 1621 AGCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1680

RESULT 6
 ABT23241
 ID ABT23241 standard; DNA; 1905 BP.
 XX
 AC ABT23241;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 24.
 XX
 KW Cytochrome P450; embryo; endosperm; size; seed development;
 XX antisense inhibition; co-suppression; transformed plant; gene; ds.
 OS Glycine max.
 XX
 PN WO200299063-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002MO-US17562.
 XX
 PR 05-JUN-2001; 2001US-295921P.
 XX
 PR 28-NOV-2001; 2001US-334317P.
 XX
 PA (DUPC) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Heppard EP, Nagasawa N, Sakai H;
 XX
 DR WPI; 2003-229292/22.
 XX
 DR P-PDB; ABJ26685.
 XX
 PT New isolated nucleic acid fragment for antisense inhibition or
 PT co-suppression of a cytochrome P450 polypeptide associated with
 PT controlling embryo/endosperm size during seed development in plants
 XX
 PS Claim 38; Page 90; 137pp; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment comprising a
 CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 CC 12021 base pairs fully defined in the specification, or their similar and
 CC functionally equivalent fragment or subfragment. The nucleic acid encodes
 CC a cytochrome P450 polypeptide associated with controlling embryo/
 CC endosperm size during seed development. The nucleic acid fragment is
 CC useful in antisense inhibition or co-suppression of a cytochrome P450
 CC polypeptide associated with controlling embryo/endosperm size during seed
 CC development in a transformed plant. This polynucleotide sequence
 CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.
 CC
 SQ Sequence 1905 BP; 488 A; 498 C; 463 G; 456 T; 0 other;
 XX
 Query Match 36.8%; Score 585.8; DB 25; Length 1905;
 Best Local Similarity 66.1%; Pred. No. 3.6e-188;
 Matches 934; Conservative 0; Mismatches 452; Indels 27; Gaps 5;

QY 193 TACAAACCGGAACGTTATTCCTCCGCTCCAAAAGCTTCCCTTGGTGGAGCATGTCA 252
 Db 294 TCCAAACACCTTAAATGATTCCTCCGCTCCAAAAGGCTATCTTTCAATGGAGCATGAGC 353
 QY 253 CTCATGTCAAGCACTAGCTCACCAGCAATGCTGATGACGTGAGAAATTCGAGCC 312
 Db 354 CTCATGACATCC---CTTGACACCAACCGATTTGCTGCGCTGCTCTAACCATGCAAGCC 410
 QY 313 AAGAGGTCAATGCTTTCAGCTTGAAGAGACTGCGGTGATGCTCACTGCAATCCGAC 372
 Db 411 ACCAGGCTCATGCGCTTCTCCATGAGGTGACACCGGTGATGCTGTCACGTCACCCGCCAC 470
 QY 373 GTAGCGAAAGATTCGAAATAGCCGCTTTTGTGATGACGCGGTGAAAGATGAGCT 432
 Db 471 GTGCGCAAGAGATTCCTTAACAGCTCCGCTTCCGATGTCTCCATTAAGAGATGAGC 530
 QY 433 TACTCACTGATGTTTAAAGAGCAATGCTTTTGCACACAGCTGTTTACTGCGAGACG 492
 Db 531 TACAGCTCATGTTTCAACCGCCCATCGCTTTCCTTACGCGCTTACTGCGCACCC 590
 QY 493 CTTCGCGCTATGCTTCCGAACCATCTCTTATGTAACAAACCAATCAAGAGCCGAGACG 552
 Db 591 CTCGCGCGCATGCGCGCCACAGCACTCTTCTGCCCCAACAATCAAGGCTTGGAGCTC 650
 QY 553 CAACGACGATGATCTCCAAAGCAGATGCTTGAAGTTCTTGAACCAAGAGTAAAGCA 612
 Db 651 CAGCGCGCCGAATCGCGCCGACAGATGACCACTCGTCCGAAAC---GCGCGGCG 704
 QY 613 CCCCTGTTTGTGTGATGCTTAAACCGGCTCCGCTTAAACAACATGATGCTCTGTA 672
 Db 705 GGTTCGGAATCGCAGCGCTTCTCAAGAGAGCGTCCGTCAACAACATGATGCTGAG 764
 QY 673 TTGCAACAGATATGATGCTTGAACCAACCA-----TGTTAGTTTACCTGAATG 723
 Db 765 TTGGACAAAGATATGACTTGAAGAGCAAAACATTCAGTGAACGATTAATCCGGTTA 824
 QY 724 GTGCAAGAGATATGATTTTGTGCGAAGCTTGAATGCACTGATCACTTCTTGGCTA 783
 Db 825 GTGCAACAGATATGATTTGTTGGGTACCTTAAATGGGAGAACATATCCCTTTTCTG 884
 QY 784 TCGGAGTTTATCTCTCAAGACTCCGCTAGATGTTCACTCGTACCAAGTAAAC 843
 Db 885 AAGACCTTGAACCTTCAAAAATCCGTTTACTGCTCAAACTCGTCCCGCAAGTAAAC 944
 QY 844 CGGTTGTATCCCGATTAATCCGAACACCGTAAATCAACCGGTATTTGCTCGTGAAT 903
 Db 945 CGGTTGTATCCCGATTAATCCGAACACCGTAAATCAACCGGTATTTGCTCGTGAAT 1004
 QY 904 TTGCTGACGTTTGTCTCTCCATGATGCTTGAATTAATTCGACCGCGGACATATC 963
 Db 1005 TTGCTGATGTTTGTCTCTCTCCAGATGCTCCGATTAATTTGCTCACTCGACATGAT 1064
 QY 964 GCCGTTTGGAGATGATATTCAGAGAAACAGACACAGTGGGCTTTAATCGAGTGG 1023
 Db 1065 GCTGTCTCTGGAAATGATATTTAGGGGACGACACGCTGGGGTTTGAATGAGTGG 1124
 QY 1024 ATCTCTGCTAGATGATGCTTCAATCCAGATATGATTAACAACGATGACAAACGAGCTGAGT 1083
 Db 1125 ATTAATGCAAGATGATGCTTCAATCCAGATATGATTAACAACGATGACAAACGAGCTGAGT 1184
 QY 1084 CAAATAGT---CGGAAATCAAGAGCCCTAGATGATCTGACTTGGCTTCACTTCATAT 1140
 Db 1185 GCGGTGTTGAGAGTGTGGCGCGCTTTGAAGAGAGACGTTGGCGGACGCGGTAT 1244
 QY 1141 CTAACGCTGTGTGAGAAAGATTTGAGGTTTATCTTCAAGCCCACTTATCATGG 1200
 Db 1245 CTTCGCGGTGTGAGAGAGATTTGAGGTTTATCTTCAAGCCCACTTATCATGG 1304
 QY 1201 GCCGTTTGGCATTAACAGACAGATGTTGATGTTGCTTGGCGAGGAGACACA 1260
 Db 1305 GCCGCTTGGCATTAACAGACAGATGTTGATGTTGCTTGGCGAGGAGACACA 1364
 QY 1261 GCAATGTTGAACATGTGGGCGCTTATCCGATGATCAACGTTGGGTTGATCTTTGGAG 1320

Db 1365 GCCATGTTATATGTGGCCATAGAGAGGACCCGAGAGTGTGGCTGACCACTGAT 1424
 QY 1321 TTAAACCTGAGAGCTTGTGGCCAAAGAGTGAAGTGAATTTTGTCTTGGGTG 1380
 Db 1425 TTCAAGCCCGAGAGGTTCAATG-----GGCTGAGAGCGAGATTTCTGTCTCGGCTG 1478
 QY 1381 GATTGAGACTTGCACCTTTCGGGTGCGGTCTGCGATTTTGGCCCGGAGAAATCTTGT 1440
 Db 1479 GATCTGAGGCTGCTCAATTCGAGTCCGGTGAAGAACTGCCCCGAGAAATCTTGGT 1538
 QY 1441 TTATCAACCTGATGTTTGGACGCGATGATGTTACATAGATGATGATGAGACGCTCC 1500
 Db 1539 TTGAGCAACCGTGACTTCTGGGTGCGAGGCTTTTGACAGATTTGAATGCTACATCT 1598
 QY 1501 GATGTTAAGCGGCTTGAATCTTATCTGAGAACTGAGGCTTCTTGCAGATGCTTATCT 1560
 Db 1599 GATGAGGGAAGTTGATCTTACAGAGGTCTGAGGCTCTGCTGTGAATGCTTACCCG 1658
 QY 1561 CTTCCTGCTAATTTGCGCGGTAGCGCAGATTAA 1593
 Db 1659 CTCTATGTTAAAGTTCCGCTTACGCGGTGATTA 1691
 RESULT 7
 AAX55600
 ID AAX55600 standard; DNA; 1799 BP.
 XX
 AC AAX55600;
 XX
 DT 26-JUL-1999 (first entry)
 XX
 DE Clone 040 ovule-specific gene encoding a cytochrome P450 monooxygenase.
 XX
 KW Ovule-specific gene; genetic engineering; ovule; nutritional value;
 XX fruit; seed; crop plant; cytochrome P450 monooxygenase; ss.
 OS Phalaenopsis sp.
 XX
 PN US5907082-A.
 XX
 PD 25-MAY-1999.
 XX
 PP 17-NOV-1995; 95US-0560398.
 XX
 PR 17-NOV-1995; 95US-0560398.
 XX
 PA (REGC) UNIV CALIFORNIA.
 XX
 PI Nadeau J, O'Neill S;
 XX
 DR MPI, 1999-337265/28.
 XX
 DR P-PsDB; AAT13438.
 XX
 PT Ovule-specific gene expression - useful for targetting gene
 XX expression to ovules or for modulating ovule development
 XX
 PS Examples; Columns 33-36; 32pp; English.
 XX
 CC The invention provides ovule-specific genes which are useful in
 CC genetically engineering plants. The use of specific promoters from the
 CC genes of the invention allows directed expression (or inhibition) of
 CC desired heterologous genes in ovules e.g. to increase the nutritional
 CC value of fruit or seed, or to inhibit the development of ovules. The
 CC polynucleotide sequences (AAX55599, AAX55601-604) are expressed uniquely
 CC during ovule development and as such they play a role in regulating
 CC development and establishing the specialised identities of cells in the
 CC ovule. The isolated genes or their promoter sequences are able to be
 CC used to produce improved crop plants. The present sequence represents a
 CC clone 040 ovule-specific gene encoding a cytochrome P450 monooxygenase
 CC from the pollen tubes of Phalaenopsis.
 XX
 SO Sequence 1799 BP; 456 A; 362 C; 428 G; 548 T; 5 other:

Query Match 22.3%; Score 355; DB 20; Length 1799;
 Best Local Similarity 58.0%; Pred. No. 1,6e-109;
 Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTGGAGCCCAAGAGCTCAGCTTTGAGCTTTGAGAGAGACTGCGTGATCGTCACAGTC 363
 DB 273 TTGACGCGCTCGCGCTCATGCGCTTCTCGGTGGGCTCACTGCTTCATCGTTCAAC 332
 QY 364 AATCCGAGTAGGAGAAAGATCTGAATAGCCCGGTTTGTGATGAGCCGTTAA 423
 DB 333 CACCCGAAACCGCAAAAGAGATCTTCAAGCCGAGCCTTGCTGATGGCCCATTA 392
 QY 424 GAATCGGCTTACTCATGATGTTTACAGAGCAATGCTTTGACACACAGGTTTAC 483
 DB 393 GAATCAGCATACGAACTTCTGTTTAACTGCGCTATGGGTTTTCGCCATTTGGGATAC 452
 QY 484 TTGGGAAAGCTTGGCGGATCGCTTGAACCATCTCTTTAGTCAAAAACAATCAGAGA 543
 DB 453 TGGAGAACTGAGAGAGATTTGCTCCATATCTTTTCACTCGCGCGAGTTGATCG 512
 QY 544 GCCGAGCGCAACGACAGATGATCTCAAGCCAGTGTGAGTTTCTTGAAGAAACA 601
 DB 513 TTGAGAGAGCAACGAGATGAGATTTGGGAGAAATGCTGGGAGTATGAAAGAAATG 572
 QY 602 -GTAGTACGAAACCTGTTTGTTCGAGATGCTTAAACGGGCTGCTTAAACAATG 660
 DB 573 GAGAGAAATGAGATGATGAGAGAGAAATGTCACGAGCTTTGAAATACATC 632
 QY 661 ATGTGCTCTGATTTGCAACAAGATGAGCTTTGAAAAAACATGTTGAGTACGTAA 720
 DB 633 ATGTGATGTTTGTGGAAAAAGTTGATTTGCAAAAGATGAGGGGTTGAGCTTGA 692
 QY 721 ATGCTC--GAAAGAGTTATGATTTGCTCGAAAGTTGATGATGATGATGATGAT 777
 DB 693 TTGATCCTTAAAGAGATGAGATGATGATGATGATGATGATGATGATGATGAT 752
 QY 778 TGGCTATCGAGATTTGATCTCAAGACTCCGCTGATGATGATGATGATGATGAT 837
 DB 753 CTTTGGAGATGTTGATTTGCAAGATGATGAGAGAGATGATGATGATGATGATG 812
 QY 838 GTAACCGGTTTGTATCCCGATTTATCCGAAACACCGTATCAAC-- 885
 DB 813 GTCAATGATTTGTGAGAGATGATGATGATGATGATGATGATGATGATGATG 872
 QY 886 -----GGTATTTGCTGATGATTTGCTGATGATGATTTGCTGCTCCCTCATGAT 939
 DB 873 ATTGATGAGGAGAGAGATGATGATTTGATGATGATGATGATGATGATGATGAT 932
 QY 940 AATATTATCGACCCGAGCATTAATGCGCGCTTTTGGAGATGATTTCAAGAGAAACAG 999
 DB 933 AGACTCTCAGATCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 992
 QY 1000 AAGATTGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1059
 DB 993 ACTGTTGCACTCTATTTGATGATGATGATGATGATGATGATGATGATGATGAT 1052
 QY 1060 TCAACGATCAAAAACGAGCTGATGATCAAGTATGCGGAAATCAAGAGCCCTGATGAT 1119
 DB 1053 TCGAAGGCAACAAGTTGATGATGATGATGATGATGATGATGATGATGATGATG 1112
 QY 1120 GACTTGGCTTCACTTCAATGATCAAGCTGATGATGATGATGATGATGATGATGAT 1179
 DB 1113 GATATCAACGATTTCTTATCTCAATCTATGATGATGATGATGATGATGATGATG 1172
 QY 1180 CCAAGGCCACTTCTATCATGAGGCGGTTTGGCCATCAACAGACAGATGATGATGAT 1239
 DB 1173 CCGGGGCTTATTTGATGAGGCTCGCTAGTATCAAGAGATGATGATGATGATGAT 1232
 QY 1240 CTTGTTCCGAGAGAGCAACAGATGATGATGATGATGATGATGATGATGATGATG 1299
 DB 1233 ATGATTCCTGCTGGAGAGATGATGATGATGATGATGATGATGATGATGATGATG 1292

QY 1300 GTTGGGTTGATCTTTGAGTTTAAACCTGAGAGGTTGCGCAAAAGAGTGAAGTG 1359
 DB 1293 AACTGGGCTGAGCTTAAACAATTCAATCTGATGATGATGATGATGATGATGATG 1344
 QY 1360 GAGTTTTCGTTCTTGGGCTGAGATTTTGAAGCTTGCACCTTTCGGGCTGCTGAT 1419
 DB 1345 ----TCAATATCTTGGTTCGATTTAAGTTGACACCTTTCGGGCTGCTGATGAT 1400
 QY 1420 TGGCCCGGAGAAATCTTGTGTTTACATGCTTATGTTTGAACGGGATGATGATGAT 1479
 DB 1401 TGGCCCGGAGAAACGATGAGATGAGCTGAGCTGATCTTGTGTTGCTGATGATG 1460
 QY 1480 GAGTTTGAATGAGGAGCGCTCCGATGATGATGATGATGATGATGATGATGATG 1539
 DB 1461 AGCTTCAATGCTTCTCTC--GAGAAATGATGATGATGATGATGATGATGATG 1517
 QY 1540 TCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1562
 DB 1518 TCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1540

RESULT 8
 ABT23246
 ID ABT23246 standard; DNA; 1930 BP.
 XX
 AC ABT23246;
 XX
 DT 01-MAY-2003 (first entry)
 XX
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 34.
 XX
 XX Cytochrome P450; embryo; endosperm; size; seed development;
 XX antisense inhibition; co-suppression; transformed plant; gene; ds.
 XX
 OS Aquilegia vulgaris.
 XX
 PN WO200299063-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US17562.
 XX
 PR 05-JUN-2001; 2001US-295921P.
 XX
 PR 28-NOV-2001; 2001US-334317P.
 XX
 PA (DUPO) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Heppard BP, Nagasawa N, Sakai H;
 XX
 DR WPI; 2003-229292/22.
 DR P-PSDB; AB226650.
 XX
 PT New isolated nucleic acid fragment for antisense inhibition or
 PT co-suppression of a cytochrome P450 polypeptide associated with
 PT controlling embryo/endosperm size during seed development in plants
 XX
 PS Claim 38; Page 97-98; 137p; English.
 XX
 CC The invention relates to an isolated nucleic acid fragment comprising a
 CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 CC 17201 base pairs fully defined in the specification or their similar and
 CC functionally equivalent fragment or subfragment. The nucleic acid encodes
 CC a cytochrome P450 polypeptide associated with controlling embryo/
 CC endosperm size during seed development. The nucleic acid fragment is
 CC useful in antisense inhibition or co-suppression of a cytochrome P450
 CC polypeptide associated with controlling embryo/endosperm size during seed
 CC development in a transformed plant. This polynucleotide sequence
 CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.
 SQ Sequence 1930 BP; 502 A; 320 C; 433 G; 675 T; 0 other;

Query Match 22.1%; Score 352; DB 25; Length 1930;

Best Local Similarity 56.3%; Pred. No. 1.8e-108;
Matches 786; Conservative 0; Mismatches 563; Indels 45; Gaps 5;

```

QY 197 AAACCGGAAACGTTATCCCGGTCGAAAAGGCTCCCTTGGTTGGAGCATGTCATCA 256
DB 213 AAGCTAAAACCGCATTCCTGACACATCGGGTTCTTTATGGGTTGGCTTGGCTT 272
QY 257 TGTCAAGCACTAGTCAACGACGATGCTGATGCAATGCAATGCAATGCAATGCA 316
DB 273 TTACTGCTTACTACTACATAGATTGCAATCTTCTAAACCTTTAAAGCTATT 332
QY 317 GGCTCATGCTTACGCTTAGAGAGACTGCGTATCGTCACTGCAATCCCGACGTAG 376
DB 333 CTTTAAAGGCTTTTCTGTTGGTTTACTCGTTTATCATATCAAGTTGCTGATAG 392
QY 377 CGAAAGAGATTCGAATAGCCGGTTTGGTGTGATGACCGGTTAAAGATGGCTTACT 436
DB 393 CAAAAGAGATTCCTTAATAGTTCTTCTTGGTATGACCTGTTAAGAAATGCTTAG 452
QY 437 CACTGATGTTTACAGAGCAATGTTTGGACCAACGCGTTTACTGGCGAAGCTTC 496
DB 453 AACTTTGTTTCAAGCAATGCGTTTGGCTCTTTGGTGAATTTGGAAGATCTGA 512
QY 497 GCGGATCGCTTGAAACCATCTCTTATGACAAAACAATCAGAAAGCCGAGAGCA 556
DB 513 GAAGATCTCAGTACCCATTTTATCAGTCAAAAGAAATTAACCGGTTTGGTACAT 572
QY 557 GAGGATGATCTCAAGCCAGATGTTGATGTTCTTGAATAACAGAGTAG---TAAGAAC 613
DB 573 GAAGTGAATAGAGAAAATATTAATGAGATTAAATGCAATGGGTTAAATGGGG 632
QY 614 CCTGTTTGTTCGTGAGTTGCTTAAACGCGCTGCTTAAACAATGATGCTCTGAT 673
DB 633 AAGTTGAAGTTAAAGGGTATACACTTGGGCTTTAAACAATGATGATGAGGTTT 692
QY 674 TCGGACAGAGTATGAGCTTGAAATAAACATGTAGTTACGT-----GAAATGG 724
DB 693 TTGGACGTTTATGATTTTAAACAATTAATGATGATGAGGTTTAAACTTGAAGATT 752
QY 725 TCGAAGAGGTTATGATTTGCTCGAAGCTTGAATGACATGATCACTTCCCTTGG 784
DB 753 TGAATGAGGGATGATGCTTGGGATTTTAACTGAGATGATCACTTCCCTTATGG 812
QY 785 CGGAGTTTATCCTCAAAAGCTCCGCTTATGATGTTCCACATCTGATCAAAAGTTAAC 844
DB 813 GCTGTTGATTTTGCAGAGATGAGAAAGAAAGAGAGAGTGTGTTCTAAGGTGAATA 872
QY 845 GATTGTTATCCCGAATTAATCCGAAACACCGTATCAACCGGT-----888
DB 873 TTTTGTGAAAAATTTATGAAAGACAGAAACAGAGATTAATGAGTTTGGGTC 932
QY 889 -GATTGCTCGTATTTGCTGAGCTTTTGTCTCTCCCTCAATGGTCAATTAATTA 946
DB 933 AAGATGTTGTTGGATCTTGTGATGTTCTTGTGATTTGGAAGAAACATAGTCTCA 992
QY 947 CCGACCGGACATTAATCGCGTTCTTGGAGATGATTAATCAAGGAAACAGACAGTTG 1006
DB 993 GTGACTGTGACATGATGCTGTTCTTTGGAAATGATCTTTAAGGGCAGACAGCTAG 1052
QY 1007 CGGCTTAAATGAGTGAATCTCGTGAAGTGTCTTCAATCAAGATGCAATCAACGG 1066
DB 1053 CAATCTCTTAAGATGATCTTGCAGAAATGCGCCCTCAATCAAGATTAATCAAGCA 1112
QY 1067 TACAAAACGAGCTGATCAAGTAGTGGGAAATCAAGAGCCATAGATGATGACTG 1126
DB 1113 CCCATCTGAATTAATGACATGCTGTGACCTAAATGCAATGATGATGATGATGATG 1172
QY 1127 CTTCACTTCATATCTAATCGCTGTGTGTAAGAAAGATTAAGGCTTCAATCTCAAGCC 1186
DB 1173 CCAACCTTCTTATCTCAAGAGATGAGAAAGAAATCTTAAGGGTCAACCTCTGAGCC 1232
QY 1187 CACTTATATATGAGGCGGTTTGGCCATTAACAGACAGATGCTTGAATGCTGCTTGT 1246

```

```

DB 1233 CCCTCTTGTGTGGGACGACATGATTCATGATGTCATATTTGGGAAGAACTTATCC 1292
QY 1247 CGGAGGAGACCAACGAAATGATGAACATGNGGCCCGATGCGATGATCCACAGTGGG 1306
DB 1293 CAGCTGGACCTACTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1352
QY 1307 TTGATCTTTGAGATTAACTGAGAGTGTGTCGAAAGAAAGATGAGTGAAGTTT 1366
DB 1353 CTGAGCAATTAATTAACCCGAGATTCATGATGAAGATG-----TGA 1400
QY 1367 CGGTTCTTGGGTGCGATTTGAGACTTCCACTTTCCGTTGGGTGCGTGGATTTGCC 1426
DB 1401 GCATTATGAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1460
QY 1427 GGAAGAACTTGTGTTTACTACCGTTATGTTTGAACGCGAGATGATGATGATGATG 1486
DB 1461 GAAGGCTTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1520
QY 1487 AATGAGGACGCTCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1546
DB 1521 AATGAGGCTTCTTGAAGGT---GATGATGATGATGATGATGATGATGATGATG 1577
QY 1547 AGATGCTTAATCCCT 1562
DB 1578 AATGAGAAAGCCCTT 1593

RESULT 9
ID ABT23240
ID ABT23240 standard; DNA; 1934 BP.
XX
AC ABT23240;
XX
DT 01-MAY-2003 (first entry)
XX
DE Seed development embryo/endosperm size alteration DNA SEQ ID 22.
XX
KM Cytochrome P450; embryo; endosperm; size; seed development;
XX antisense inhibition; co-suppression; transformed plant; gene; ds.
XX
OS Glycine max.
XX
PN WO20029063-A2.
XX
PD 12-DEC-2002.
XX
PF 04-JUN-2002; 2002WO-US17562.
XX
PR 05-JUN-2001; 2001US-295921P.
XX
PR 28-NOV-2001; 2001US-334317P.
XX
PA (DUPO ) DU POINT DE MEMOIRS & CO E I.
XX
PI Cahoon RE, Heppard EP, Nagasawa N, Sakai H;
XX
DR WPI: 2003-229292/22.
XX
P-PSDB; A3026684.
XX
PT New isolated nucleic acid fragment for antisense inhibition or
XX co-suppression of a cytochrome P450 polypeptide associated with
XX controlling embryo/endosperm size during seed development in plants
XX
PS Claim 38; Page 88-89; 137pb; English.
XX
CC The invention relates to an isolated nucleic acid fragment comprising a
CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
CC 17201 base pairs fully defined in the specification, or their similar
CC functionally equivalent fragment or subfragment. The nucleic acid encodes
CC a cytochrome P450 polypeptide associated with controlling embryo/
CC endosperm size during seed development. The nucleic acid fragment is
CC useful in antisense inhibition or co-suppression of a cytochrome P450
CC polypeptide associated with controlling embryo/endosperm size during seed
CC development in a transformed plant. This polynucleotide sequence

```

CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.

XX Sequence 1934 BP; 412 A; 386 C; 547 G; 589 T; 0 other;

Query Match 19.9%; Score 316.6; DB 25; Length 1934;
 Best Local Similarity 55.1%; Pred. No. 2.1e-96;
 Matches 750; Conservative 0; Mismatches 569; Indels 42; Gaps 5;

254 TCATGTCAAGACCTTACCTCAACGAGATGCTGATGACGCTGAAGAAATTGGAGCCA 313
 255 TCTTATCTGCTCCACGCTCAGCTGCTTATCAAACTGCGCTTATTAACCAACCGG 314
 314 AGAGCTCATGAGCTTTGAGTGAAGAGACGCGGATGCTGACGTAATCCGAGC 373
 315 AAAAATCATGAGCTTTCCATCGCTTAAACCGCTTCTCATCTCCAGGAAACCGAGA 374
 374 TAGCGAAAGATTCGAAATGACCGGCTTTTGTGATGACGCGGTTAAAGATCGGCTT 433
 375 CCGCTAAGAGATTCCTCGGACGCCAGCTTCCGCTGATAGGCGCGTGAAGAAATCCGCT 434
 434 ACTCATGATGTTTACAGCAATGCTTTTGAACCAACGCTTTTACGCGGAGAGC 493
 435 ATGAGCTTCTCTCCACCGGAAAGGCTTTGACCGGATGAGGAGTCTGAGGAATT 494
 494 TTCGCGCTATGCTTCGAAACCATCTCTTATGATCAAAACAAATGAGAGCGGAGACG 553
 495 TGAGGAGATCTCAACCTTACATCTCTCCCGGAGAGATCAACGCGCTCTGAATCCT 554
 554 AACGACGATGATCAACGAGATGCTGAGTTCCTGAAAACAGAGTATG---AACG 610
 555 TCAGAGAGGAGTGAATTAAGATTAAGTTGAACAGTTAAACCAATGAGAGAACCC 614
 611 AACCTGTTTGTGCTGAGTGTCTTAAACGCGCTGCTTACAGCATGATGTCTCTG 670
 615 AACATGTTGAGTTAAGAAATTTCACTTATGTTGTTGAACATGATGATGAGAGC 674
 671 TATTGGAACAGATGATGAGCTT---GAAAAAACATGTTGATGATGATGATGATG 727
 675 TGTGTTGATGCTTATGATGATTTAAGAGGAGGTTGAGCTTGAAGGTTTGGTGA 734
 728 AAGAGTTATGATTTGCTCGGAAAGTTGAATGAGTATGATCACTTCTGCTATCGG 787
 735 GTGAAGGATGATGATGTTGAGGTTTAACTGAGAGCACTTTCCGCTTTTGGGGT 794
 788 AGTTGATCTCAAGACTCCGCTGATGATGTTCCACCTGTACCAAAAGGTAACCGGT 847
 795 GGTGAGATTTGAGGAGTGAAGAGAGGATGATGATGATGATGATGATGATGATG 854
 848 TTGATCCCGGATTTATCCGACACCGTAAATCAACCGG----- 887
 855 TTGTTGAGAGGTTTAAAGAGCAATGAGGTAAGAGGAGAGGAGGATGATGATGAGG 914
 888 -TGATTTGCTGATTTGCTGACGTTTGTCTCTCCCTTCATGATGATGATGATGAT 946
 915 ATGAGAGAACTGGGATTTTGTGATGTTTGTGATGATGATGATGATGATGATGAT 974
 947 CCGACCCCGGATTAATCCGCTTCTTGGGAGATGATTAATGAGAGAACAGACAGTTG 1006
 975 GTGAAGCTGATGATGATGCTGTTCTTTGGGAAATATATTTAAGGGAACCTACAGG 1034
 1007 CGGCTTAAATGAGTGAATGCTCGTGAAGATGCTTCTTCATCAAGATGATGATCAACG 1066
 1035 CAATTCTGTAGAGTGAATGCTCGCTCGATGATGCTTCCACCTTAAATTCAGAGAA 1094
 1067 TCAAAAGAGGATGATCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 1126
 1095 CACAGGCGAATTAATCTGTTTGGGATCTTCCAGGCGCTTATCCAGAGCAATCTTC 1134
 1127 CTTCATCTTCAATATCAAGGCTGAGTGAAGAAAGTATGAGGCTTATCTTCAGGCGC 1186
 1155 CCAAGCTGCTACCTTCAAGTGAATGATTAAGAAACCTTGTGATGATGATGATGATG 1214

QY 1187 CACTTCTATCATTGAGGCGGCTTTGGCCATACAGACAGCATGTTGATG---TCGCTTGG 1243
 DB 1215 CCTTACTCTCTGAGGCTGCTGCTTGTGACAGACGTTACCTGAGCGGCAAGCATGGA 1274
 QY 1244 TTCCGAGAGGACCAACAGCAATGTTGAACATGTTGGCCGTTATCGCATGATCAACGTTG 1303
 DB 1275 TTCCCAAGGACCAACCGCATGTTGAACATGTTGGCCATACCAACAGAGAGGTTG 1334
 QY 1304 GGGTTGATCTTTGAGTTTAACTGAGAGGTTGTTGAGCAAAAGAGTGAAGTGAAGT 1363
 DB 1335 GGGCGAGCGCGAAGATTTAGCCCGAGCGGTTGTG-----GAGAGAGATG 1382
 QY 1364 TTTCGCTTCTTGGGTCGATTTGAGATCTGACCTTTGCGGTCGGTCTGATTTGCC 1423
 DB 1383 TGAGCATATAGGAGTCTGATTTGAGGTTGGCACCTTTGCGGTCGAGAAAGATGTGCC 1442
 QY 1424 CCGGAAAGAAATTTGTTTACTACGTTATGTTTGAAGGCGATGATGATGATGATGAT 1483
 DB 1443 CTGGGAAGGCTCTGTTGCTGCTGTTATCTTGTGCTGCTGATGCTGATGCTTCAAA 1502
 QY 1484 TTGAATGAGGACCGTCCGATGTTAAGGCGGTTGACTTATCTGAGAACTGAGGCTTTCT 1543
 DB 1503 TTCAATTGGCTTCACTGATGATGTTGTTCTGTTGAGTGAATGATGATGATGATGAT 1562
 QY 1544 GCGAGATGCTATATCTCTCTCTCTTAATTTGCGCGGTTAGG 1584
 DB 1563 TGAGATGAAAGAGCACTGTCTTGAAGGCTGCTGCTGAG 1603

RESULT 10

ABT23249
 ID ABT23249 standard; DNA; 1687 BP.

XX ABT23249;
 DT 01-MAY-2003 (first entry)
 XX

DE Seed development embryo/endosperm size alteration DNA SEQ ID 40.

XX Cytochrome P450; embryo; endosperm; size; seed development;
 KM antisense inhibition; co-suppression; transformed plant; gene; ds.

XX *Astroemeria caryophylla*.

XX W0200299063-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002WO-US17562.

XX 05-JUN-2001; 2001US-295921P.

XX 28-NOV-2001; 2001US-334317P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

XX Cahoon RE, Heppard BP, Nagasawa N, Sakai H;

XX WPI; 2003-229292/22.

XX P-PSDB; ABJ2693.

XX New isolated nucleic acid fragment for antisense inhibition or

XX co-suppression of a cytochrome P450 polypeptide associated with

XX controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 101-102; 137pp; English.

XX The invention relates to an isolated nucleic acid fragment comprising a

XX promoter consisting essentially of a sequence having 2000, 2022, 8300 or

XX 11201 base pairs fully defined in the specification, or their similar and

XX functionally equivalent fragment or subfragment. The nucleic acid encodes

XX a cytochrome P450 polypeptide associated with controlling embryo/

XX endosperm size during seed development. The nucleic acid fragment is

XX useful in antisense inhibition or co-suppression of a cytochrome P450

CC polypeptide associated with controlling embryo/endosperm size during seed
CC development in a transformed plant. This polynucleotide sequence
CC represents a gene relating to the alteration of embryo/endosperm size
CC during seed development of the invention.

SQ Sequence 1687 BP; 246 A; 636 C; 456 G; 349 T; 0 other;

Query Match	Score	DB	Length
18.2%;	290.6;	25;	1687;

Best Local Similarity	54.18;	Pred. No. 1.4e-8/;
Matches 712; Conservative	0;	Mismatches 579; Indels 24; Gaps 5

QY	257	IGTGAAGCACTCTAGCTCAGCCGACGGAATCGCGTATGCAGCTGAGAAATTCGAGCGCAAGA	316
Db	193	TCCTCGGCCCCCTCGGCGCAACCGCTCCCTCGCGCCGCTCTCTCTCCCTCGCGCCCTCC	252
QY	317	GAGCTCATGGCTTTCAGCTTAGAGAGACTCGCGTATGCTGACGCTGCAATCCGACGTA	376
Db	253	CCCTCTCTCTCTCTCTCTCGGCGCTCAACCGCTCATGTCTCTCCCAACCCCTCCACCG	312
QY	377	CGAAGAATCTGAATAGCCCGGTTTTTGCTGATCGACCGGTTAAATAATCCGTTACT	436
Db	313	CGAAGAATCTCTCCAGCTCCGCTTCGCGGACCGGCCCATCAAGACTCCGCTTACG	372
QY	437	CAGTATGTTTAAACAGACCAATTGGTTTTTGACACACACGAGTWTACTGAGGAACGCTTC	496
Db	373	GCTCTCTTTCACCGCGGCAATGGCTTCGCCCCCTTGCGTGAATCACTGGGCAACTCC	432
QY	497	GCCGATCGCTTCCGAACCATCTCTTTAGTACAAACAAATCAAGAAAGCCGACGCAAC	556
Db	433	GCCGATCTCGGCAACCAACCTCTTCAGACCCCAACCGGCTCTCGGCTCCGCCCCCTCC	492
QY	557	GACGAGTATCTCAAGCCAGATGGTTGATTTCTTGAATAAACAGAGTGTATACGAACCT	616
Db	493	GCCGGAATCGGCTCCGCGCGCTCCCAAGTGGCTCTCCCTATAGCCACCAACGCGC	552
QY	617	GTTTGTTCGTGAG--TTGCTTAAACGGCGCTGCTTAAACAATGATGTGCTGTAT	673
Db	553	AGGTGAGATCAAGGCGCTCTCCCACTTCGCTCCCTCAACACACTATGCGCAGCGT	612
QY	674	TCGCAAGAGATGAGCTTGAAAAA---ACATGTGAGTACGTAAATGGTCGAAG	730
Db	613	TGCGCGCGCTCTACACATTCGCCACCGGGAGCGGCTGAGTCAAGCCTTGGTCAAGC	672
QY	731	AAGTTATGATTTGCTCGAAAGTTGAATTGACTGATCACTTCCTTGGCTATCGAGT	790
Db	673	AGGGTACGAGCTCTGGCGCTTCAACTGGGGGAGCACTTTCGCGTGTGGCTGGT	732
QY	791	TTGATCTCAAGAATCCGCGTCTAGATGTGCCACTCGTACCAAGTAAACCGGTTG	850
Db	733	TTGACTTCAAGGGGTCAAGCGGAGTGCAGAGCCCTGCTCAGCCGCTCAACGCTTTG	792
QY	851	TATCCCGATTATATCGAACAC---GTAATCAACCGGTGATTTGCTCGTATTTG	907
Db	793	TCGCGCGATATTCAGCGACGCCCAAGAGCGGCGACGCGCTCCGTACAGCAGCGCG	852
QY	908	TCGAGTTTGTCTCTCCCTCCATGAGTTTCAAGATAATTTACGACCCGACATATCGCGC	967
Db	853	CGTGAATCTCTCAACGCTCTGTCTCGACGAAACCTCTCGATTTCGACATGATGGCG	912
QY	968	TTCTTTGGAGATGATATTCAGAGGAACAGACAGTTGCGGCTTTAATCGAGTGAATCC	1027
Db	913	TTCTCTGGGAAGTATCTTTGCGGGAAGGATAGGTGGCCATCTGCTGAGTGATCA	972
QY	1028	TCGCTAGAGTGTCTTCAATCCAGATATGCATCAACGTTACAAACGAGCTGATCAAG	1083
Db	973	TGGCAGAGATGTGTGTGACACCGGAATTCAGGCCAAAGCTCAACCGAGATCGACGCG	1033
QY	1088	TAGTGGGAAATCAAGAGCCCTAGATGATCGACTTGAGCTTCACTTCATATCAACGG	1147
Db	1033	TTGTGGGCGGTAGAGGTGAGTGGCGAGACGCGAAGTGGCAACTCTTACCTTCAAG	1092
QY	1148	CTGTGATGAAGAGATTTAGAGCTTCACTCTCGAGCCCACTTTCATATCAGGCGCGTT	1207

Db 1093 CATTGCTCAAGAAATGCGCTGAGATGACACCCCCCGCGCTGCTCTCTGGGCTCGCC 1152

Qy 1208 TGGCCATPACAGACACAGATGTTGATGTGTGTTCTGTTCCGGCAGGGAACACAGCAATGG 1267

Db 1153 TCCGAGTCCATGAGTGTCACGTCCGGGGGCCACTTCGTCGGCGCGACACACCGCATGG 1212

Qy 1268 TGAACATGTGGGGCCGTATCCGATGATCACACGATGAGGTGATCTTTGAGTTTAAC 1327

Db 1213 TGAACATGTGGGCCATATCGCACGCGAAATCTGGCCGAGCCGAGGTGTTCAAC 1272

Qy 1328 CTGACAGATTCTGTGGCAAAAGAAAGGTAGGTGAGTTTCGGTCTTGGGTGGATTGTA 1387

Db 1273 CGGAGAGTTTGTC-----GACACAGATGTAGACATCTTGGGCTCGGATCTCC 1320

Qy 1388 GACTTGACCTTTCCGTCGGGTCGTGCGATTTCGCCCGGAGAAATCTTGGTTTACTA 1447

Db 1321 GGCTGGCGCGCTCGGTCGGGAGAGAGGTGTGTCCCGCAAGCGCATGGGCTGGCCA 1380

Qy 1448 CCGTTATGTTTGGACGGCGCATGATGTTAATAGATTGAAATGGGACACGTCGATGGTA 1507

Db 1381 CCGGCGATCTCTGGCTGGCTCAAGCTGTTCAAGACTTCAAGTGGTGGCTTCCGA---CA 1457

Qy 1508 ACGGCGTTGACTATCTGAGAAATGAGGCTTTCTTGCGAGATGGCTAATCTCT 1562

Db 1438 ATGGGTTGATCTCTCGGAAATCTGAAAGATGTCTCTGAGATGAAGTCCCTCT 1492

RESULT 11
 ABT23284
 ID ABT23284 standard; DNA; 1545 BP.
 AC ABT23284;
 DT 01-MAY-2003 (first entry)
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 96.
 KW Cytochrome P450; embryo; endosperm; size; seed development;
 KW antisense inhibition; co-suppression; transformed plant; gene; ds.
 OS Zea mays.
 PN WO200295063-A2.
 PD 12-DEC-2002.
 PF 04-JUN-2002; 2002WO-US17562.
 PR 05-JUN-2001; 2001US-295921P.
 PR 28-NOV-2001; 2001US-334317P.
 RA (DUPO) DU PONT DE NEMOURS & CO E I.
 RA Cahoon RE, Hppard EP, Nagasawa N, Sakai H;
 DR WPI; 2003-229292/22.
 DR P-PEDB; ABJ26714.
 PT New isolated nucleic acid fragment for antisense inhibition or
 PT co-suppression of a cytochrome P450 polypeptide associated with
 PT controlling embryo/endosperm size during seed development in plants
 PS Claim 38; Page 127; 137bp; English.
 CC The invention relates to an isolated nucleic acid fragment comprising a
 CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 CC 17201 base pairs fully defined in the specification, or their similar and
 CC functionally equivalent fragment or subfragment. The nucleic acid encodes
 CC a cytochrome P450 polypeptide associated with controlling embryo/
 CC endosperm size during seed development. The nucleic acid fragment is
 CC useful in antisense inhibition or co-suppression of a cytochrome P450
 CC polypeptide associated with controlling embryo/endosperm size during seed

CC represents a gene relating to the alteration of embryo/endosperm size
 during seed development of the invention.

XX Sequence 1545 BP; 182 A; 572 C; 559 G; 232 T; 0 other;

Query Match 18.0%; Score 286; DB 25; Length 1545;
 Best Local Similarity 54.1%; Pred. No. 4,9e-86;
 Matches 692; Conservative 0; Mismatches 550; Indels 36; Gaps 4;

QY 306 CGAGCCAAAGAGCTTCATGCTTTCAGCTTGAAGAGACTGCGTGAATCGTCAAGTCA 365
 DB CGGCGCCCGCGCTGTCGCTTCTCCGTCGCTCAAGCGCTTCGTCGTAAGAGCA 314
 QY 366 TCCCGACGAGGAGAAAGATTCTGATAGCCCGGTTTTCGTCATCGACCGTTAAGA 425
 DB GCCGACACAGCGCGGAGCTCTGCGACGCGCCCTTGGCGACGCGCCCGCGAGGA 374
 QY 426 ATCGCTTACTCATGATGTTTACAGAGCAATGTTTTCACACACGCTGTTTACTG 485
 DB CGCGCGCGGCGGCTCTCTTCCACCGCGCATGCGCTTTCGCGCGGCGACTACTG 434
 QY 486 GCGAAGCTTGGCGCTATCGCTTGAACCATCTCTTATGTAACAAACAAATCAGAGAGC 545
 DB GCGCGCGCTTGGCGCATGAGCTCGCGCTTCTTCCAGCCCGCGAGCGCTGGCGCGGC 494
 QY 546 CGAGACGCAACGACGATGATCTCAAGCCAGATGATGTTTCTTGAACAAAGAGTAG 605
 DB GGGCGCGCGCGCGCGCGCATCGAGAGCGATGCGGAGACTCTCGCGCGCGCGG 554
 QY 606 TACAGAACCTCTTTCGTCGATGTTTAAAGCGCGCTTACACATGAGTGTG 665
 DB 555 ACGAAGAGTCTTCTATGCGCGCTGCTCCAGCGCGCATCTCGGACACGTCATGCG 611
 QY 666 CTCTGATTCGAGCAAGATATGATGCTTGAACAAACATGTTGATGATGAT 722
 DB 612 CACCGTTCGCGCGCGCGCTTACAGCGCGCGACCGCGAGCGCGGAGCTGAGAGAT 671
 QY 723 GGTGGAAGAGTATGATGTTGCTGGAACGTTGATGATGATGATGATGATGATGAT 782
 DB 672 GGTGGAAGAGTATGATGTTGCTGGAACGTTGATGATGATGATGATGATGATGAT 731
 QY 783 ATCGAGTTCGATTCCTCAAGACTCGGCTCTGATGTTTCCACACTGTAACAAAGTAA 842
 DB 732 CAGGTGCTGAGACTGCAAGGCGTCAAGAGGCGGTGAGAGGAGCTGAGAGTCAA 791
 QY 843 CCGGTTGATCCCGGATTAATCCGAACACCGTATCAACCGGTGATTCCTCGTGA 902
 DB 792 CGTGTTCGTCGAGATCATGAAAGACAGGCAAGAGGACGACGCGATGAGGA 851
 QY 903 T-----TTCGTCGACGTTTTCCTCCCTCAATGTTTCAGATAATT 944
 DB 852 GCGCGCGCGCGCGGAGACTTCGTCGACGCTTCGTCGAGATGAGAGGAGAGAAAGCT 911
 QY 945 ATCCGACCGGACATATGCGGCTTTTGGAGATGATTAATCAAGAGACGACAGT 1004
 DB 912 GTCGAGCTCCGACATATGCTGTCCTCTGAGAGATGATCTTTCGAGGACCGACAGT 971
 QY 1005 TCGGCTTAAATGAGTGAATCTGCTAGATGAGTTCCTTCAATCCAGATATGATCAAC 1064
 DB 972 GGGGATCTGCTGAGTGGATGAGCGGAGATGAGTGGTCAACCGCGGATCAATCA 1031
 QY 1065 GTTACAAAAGAGCTGATCAATGATCTGAGAAATCAAGACCTTATGATTAATCTGACT 1124
 DB 1032 GCGGACGCGGAGACTGAGCGCTGCTGAGGCGCGCGCGCTTTCGAGCGCGAGCT 1091
 QY 1125 GGTTCATCTCATATCTAATCGGCTGAGGAAAGAAAGATGAGGCTTCAATCCACAG 1184
 DB 1092 GGGCGCGCTGCTTACTGTCAGCGCTGTAAGAGAGAGCTCCGCTCAACCGCGG 1151
 QY 1185 CCCACTTATCATGAGCGCGCTTTCGACATPACAGACAGATGATGATGATGATGAT 1244
 DB 1152 CCGCTGCTCTGCTGAGCGCGCTGAGCGCGCTGAGCGCGAGCGGCTGAGCGCGAGTGT 1211

QY 1245 TCCGCGAGGAGACACAGCAATGATGAGCGCGCTATTCGATGATCAACGCTGTG 1304
 DB 1212 CCGCGCGGACACAGCGCATGATGATGAGCGCGCTATTCGAGCGAGCGCGCGGTG 1271
 QY 1305 GGTTCATCTTTCGAGTTCATTAACCTGAGAGCTTCGTCGCAAAAGAGTGAAGT 1364
 DB 1272 GCGGAGCGCTTCGCTTCGCGCCGAGCGGTTTCAG-----GTGAGAGAGT 1319
 QY 1365 TCGGTTTCGAGTTCGATTCGATTCGACCTTTCGAGTTCGAGTTCGATTCGCC 1424
 DB 1320 GAGCGCTGCGCGCGGAGCTCCGCTCCGCGCTTCGCGCGCGCGCGCGGTGCCC 1379
 QY 1425 CCGGAAAGATCTTGTTCATCACTGTTTTCGACGCGGATGATGATGATGAT 1484
 DB 1380 GGGCAAGACGTCGCGCTGCGCATGTCACCTTCGCTGCGCGAGCTGCGACGCTT 1439
 QY 1485 TGAATGAGGACCGTCGATGATGATGATGATGATGATGATGATGATGATGAT 1544
 DB 1440 CCGTGGCGCGCGCGCGAGCGCGCGCGCTGCACTGCGGAGCGCGCTCGCATCT 1499
 QY 1545 CGAGATGCTTAATCTCT 1562
 DB 1500 GAGATGAGAGAGCCCT 1517

RESULT 12
 ABT23283
 ID ABT23283 standard; DNA; 1758 BP.

XX ABT23283;

DT 01-MAY-2003 (first entry)

XX Seed development embryo/endosperm size alteration DNA SEQ ID 94.

XX Cytochrome P450; embryo; endosperm; size; seed development;

XX antisense inhibition; co-suppression; transformed plant; gene; ds.

XX Zea mays.

XX WO200299063-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002MO-US17562.

XX 05-JUN-2001; 2001US-295921P.

XX 28-NOV-2001; 2001US-334317P.

XX (DUP) DU PONT DE NEMOURS & CO E I.

XX Cahoon RE, Heppard BP, Nagasawa N, Sakai H;

XX WPI; 2003-228292/22.

XX P-PSDB; AB026713.

XX New isolated nucleic acid fragment for antisense inhibition or

XX co-suppression of a cytochrome P450 polypeptide associated with

XX controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 124-125; 137pp; English.

XX The invention relates to an isolated nucleic acid fragment comprising a

XX promoter consisting essentially of a sequence having 2000, 2022, 8300 or

XX 17201 base pairs fully defined in the specification, or their similar and

XX functionally equivalent fragment or subfragment. The nucleic acid encodes

XX a cytochrome P450 polypeptide associated with controlling embryo/

XX endosperm size during seed development. The nucleic acid fragment is

XX useful in antisense inhibition or co-suppression of a cytochrome P450

XX polypeptide associated with controlling embryo/endosperm size during seed

XX development in a transformed plant. This polynucleotide sequence

XX represents a gene relating to the alteration of embryo/endosperm size

XX during seed development of the invention.

XX Sequence 1758 BP; 274 A; 650 C; 545 G; 289 T; 0 other;
SQ

Query Match 17.8%; Score 283; DB 25; Length 1758;
Best Local Similarity 55.1%; Pred. No. 5,6e-85;
Matches 709; Conservative 0; Mismatches 530; Indels 48; Gaps 6;

```
QY 306 CGAGACCAAGAGCTCATGGCTTTCAGCTTAGAGAGATCCCGCATGTCAGACGTGCA 365
DB 444 CGCGGCAAGGCACTATGAGCTTCTCGGTCGGGCTGACCCGCTCTGCTGTCAGAGCA 503
QY 366 TCCGACGATAGCAAGAGATTCGATAGACCCGGTTTGTGCTGATCGACCGGTTAAGA 425
DB 504 GCGCGATAGCGCGCGAGATCTCGCCAGCCCGCGGTTGCGGAGACCGCCGTCAGAGA 563
QY 426 ATGGCTTACTCATGATGTTTAAAGAGCATGTGTTTGCACCAAGG---TGTTTA 482
DB 564 CGCGGCGCGCATCTGCTCTTCCACCGCGCATGGGCTTCGCGCTTCGAGAGCGCGA 623
QY 483 CTGCGAAGCGCTTCGCGGATCGCTTCGAAACCATCTCTTAACTAACAACCAATAGAGA 542
DB 624 CTGGCGCGGCTCCGCGGCTCCGCGCCAGACACCTGTTGCGCGCGCGCGCGCGG 683
QY 543 AGCCGAGACGCAACGACGAGTATCTCAAGCCAGATGTTGAGTTTCTTGAACAAAG 602
DB 684 TGCCGCGCACACCGCGCTCCATCGCGAGGCCATGTCGCCAGCTGCGCGCTGCCAT 743
QY 603 TAATAAGAACCTGTTTGTGTGTAGT---TGCTTAAAGCGGCTGCTTAACAACAT 659
DB 744 GCGCGCGCACGCGAGAGTCCCTCTCAAGCGCTGCTGCAATGTCATCTCTCAACACGT 803
QY 660 GATGTGCTCTGATTCGACCAAGATATGACTTTGAAAAAACCATGTTGAGTTACGTGA 719
DB 804 CATGGCCACCGTGTGGCAAGGCTACGACATGGGACCGAGAGGCGCCCTTCTGGA 863
QY 720 A---ATGCTGAAAGAGTTATGATTTGCTCGGAAGTTGAATTGACATGATCACCTTC 776
DB 864 CGAGATGTGTGCGGAGGCTACGACCTCTGGGCACTTCACTGAGCTGACCACTGCG 923
QY 777 TTGCTATCGGAGTTGATCTCTCAAGACTCCGATCTAGATGTTCCACACTGTCACAA 836
DB 924 ATTGCTAAGCATCTGACCCCGAGCGGCGGCGCGGTCGACAAAGGCTGATCCGAA 983
QY 837 GGTAAACCGGTTGTATCCGATATATCCGAACACCGTATCAACCGG-----887
DB 984 GATCGAATCGTTCGTTGCAAGATCATCTGGAGCACAGGCGCGCGCAAAATGAGG 1043
QY 888 -----TGATTGCTGCTGATTTCTGTCGAGCTTTTCTCTCCCTCCATGATTC 935
DB 1044 AGTCGTGGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1103
QY 936 AGATTAATATTCGACCCCGACATATCGCCGTTCTTTGGAGATGATATTCAGAGAAC 995
DB 1104 GGAAGACCTGTCAATCGACGACATGATCGCTGTTCTTTGGAGATGATATTCAGAGCGC 1163
QY 996 AGACACAGTTCCGCTTTAATCGAGTGAATCTCGCTAGGATGTTCTTATCATCAAGTAT 1055
DB 1164 CGACACGCTGCGATCTTGAATGATGATGATGATGATGATGATGATGATGATGAT 1223
QY 1056 GCAATCAACGCTACAAAACGAGCTGATCAAGTACGCGAATCAAGACCTTATGATGA 1115
DB 1224 CCAAGCGAAGGCCACGAGCGAGCTGACGCGCTGTGAGGACCTCGGCGCGCGTGGGA 1283
QY 1116 ATCTGACTTGGCTTCACTTCATATCTAAGCGCTGTGTTGAAAGATATATGAGGCTTCA 1175
DB 1284 CGCGAGCTGCGCAAGCTTACCTTCACTCACTGATGATGATGATGATGATGATGATG 1343
QY 1176 TCTTCAGAGCCACTTCTATCATGAGGCGCTTGGCCATTAACAGACATGTTGATGG 1235
DB 1344 CCGCGCAAGCGCGCTCTGCTGTGAGCGCGCTGCTCCACACGCGCATGTCGCGG 1403
QY 1236 TCGCTTGTTCGAGAGGAGCAACAGAAATGATGAACATGAGGCGCGATGATGATCC 1295
```

```
DB 1404 CCACCTGTCGCCGCGGACCAACAGCATGTGAACATGTGTCATCCGACGACCC 1463
QY 1296 ACACGTGTGGTATGATCTTTGAGATTAACTGAGAGGTGTGTGCAAAAGAGTGA 1355
DB 1464 CGCCATGTGGGCGGAGCGAGAAATTCGCCCGCGAGCGGTTCAG-----GA 1511
QY 1356 GGTGAGATTTTCGTTCTTGTGATGATGATTTGAACCTTTCGAGTGTGCTGTG 1415
DB 1512 GGAAGAGTGAAGCTCTCTCGGAGAGCAGCTCCGCTGAGCCCTTTCGCGCGCGCGC 1571
QY 1416 GATTGCCCCGGGAGAAATCTTGATTTACTACCTTATATGTTTGAACCGCGATGAT 1475
DB 1572 CGCCTGCCCCGGAAGTACTGCGCTGCGCACACCAACCTCTGGGTCCCAAGCTTCT 1631
QY 1476 ACATGAGTTGAATGAGGAGCCGTCGATGTTAACGCGCTGACTTATCTGAGAACTGAG 1535
DB 1632 GCACAAATTGAGTGGG-----CGCGCGCGGCGGCGTGCACCTGTGAGGCGCTGAG 1685
QY 1536 GCTTCTTGGGAGATGCTATCTCT 1562
DB 1686 CATGTGCTGAGATGCGCACGCGCT 1712
```

RESULT 13

ID ABR23237 standard; DNA; 1597 BP.

ABR23237;
01-MAY-2003 (first entry)

Seed development embryo/endosperm size alteration DNA SEQ ID 16.

Cytochrome P450; embryo; endosperm; size; seed development;
antisense inhibition; co-suppression; transformed plant; gene; ds.

Zea mays.

WO200299063-A2.

12-DEC-2002.

04-JUN-2002; 2002WO-US17562.

05-JUN-2001; 2001US-295921P.

28-NOV-2001; 2001US-334317P.

(DUPO) DU PONT DE NEMOURS & CO E I.

Cahoon RE, Hepard EP, Nagasawa N, Sakai H;

WPI: 2003-229292/22.

P-FSDB; ABR26681.

New isolated nucleic acid fragment for antisense inhibition or

controlling embryo/endosperm size during seed development in plants

Claim 38; Page 83; 137pp; English.

The invention relates to an isolated nucleic acid fragment comprising a promoter consisting essentially of a sequence having 2000, 2022, 8300 or 17201 base pairs fully defined in the specification, or their similar and functionally equivalent fragment or subfragment. The nucleic acid encodes a cytochrome P450 polypeptide associated with controlling embryo/ endosperm size during seed development. The nucleic acid fragment is useful in antisense inhibition or co-suppression of a cytochrome P450 polypeptide associated with controlling embryo/endosperm size during seed development in a transformed plant. This polynucleotide sequence represents a gene relating to the alteration of embryo/endosperm size during seed development of the invention.

Sequence 1597 BP; 279 A; 523 C; 517 G; 278 T; 0 other;

Query Match 17.5%; Score 278.2; DB 25; Length 1597;
 Best Local Similarity 54.9%; Prod. No. 2.3e-83;
 Matches 706; Conservative 0; Mismatches 533; Indels 48; Gaps 6;

QY 306 CGAGACCAAGAGCTCATGCTTTCAGCTTACGAGAGAGCTGCGGTGATGCTACCTGCA 365
 Db CGAGGAGCAAGCACTGATGCGCTTCTGCGGCTGACCGCTGCTGCTGCTGAGCA 109
 QY 366 TCCGACGTAGCGCAAGAGATTGATAGCCGCTTTCGATGCAACCGGTAAAG 425
 Db 50 CGGCGGCAAGCACTGATGCGCTTCTGCGGCTGACCGCTGCTGCTGCTGAGCA 109
 QY 110 GCGGCAAGAGCGCGCGAGATCTCGGCAACCGCGCTTGGGAGCGGCGGCAAGGA 169
 Db 426 ATCGGCTTACTGATGATGTTTAAAGAGCAATGTTTGCACACAGG--TGTTA 482
 QY 170 CGCGGCGCGCACTGCTCTTCCACAGCGCATGCGCTTGGCGCTTCCGAGAGCGGCA 229
 Db 483 CTGCGCAACCGCTTGGCGGTACGCTTGAACCATCTCTTATGACAAACAAATCAGAA 542
 QY 230 CTGCGCGGCTCGCGCTCGCTCGCGCAACCACTCTTGGCGCGCGCGCGGCGG 289
 Db 543 AGCGGAGCGCAAGAGATGATCTCAAGCGAGATGCTTGAATTTGAAAAACAGAG 602
 QY 290 TGCGGCGGACACCGCGCTCCATCGGAGAGCGCATGCTGCGCGCGCGCTGCA 349
 Db 603 TAGTACGACACCTGCTTGTGCTGAGT---TGCTTAAACGCGCTGCTTAAACAAT 659
 QY 350 GCGCGGCGCAAGAGCTGCTCTGCAAGCGCTGCTGCTGCTGCTGCTGCTGCTGCT 409
 Db 660 GATGCTCTGATTCGACAAAGATAGAGCTTGAATTTGAAAAACCATGTTAGTTAGT 717
 QY 410 CATGCGCAACCGCTTGGCAAGCGGTACGATGAGGAGCGGAGCGCGCTTCTGGA 469
 Db 718 -GAAATGCTCAAGAGATGATGATTTGCTCGAAACGTTGATGATGATGATGATGAT 716
 QY 470 CGAGATGCTGCGCGGCTGCAAGCTCTGCGGCACTGCTGCTGCTGCTGCTGCTGCT 529
 Db 777 TTGCTATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 836
 QY 530 ATTGCTCAAGATCTGCAACCGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCG 589
 Db 837 GGTAAACCGGCTTGTATCCCGATTAATCCGAACCGCTTATCAACCGG----- 887
 QY 590 GATCGAATCGTTCGTCGAAGATCATGAGAGCAAGAGCGCGCGGAAATGAGAG 649
 Db 888 -----TGATTTGCTGCTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 935
 QY 650 AGTCGTGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 709
 Db 936 AGATTAATTTATCGAAGCGCGGATATGCGCGCTTCTTGGAGATGATGATGATGATGAT 995
 QY 710 GAGAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 769
 Db 996 AAGACAGCTGCGGCTTAAATGAGAGATCTGCTGAGATGATGATGATGATGATGAT 1055
 QY 770 CGACACCGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 829
 Db 1056 GATATCAAGCTTAAACAGAGCTGATCAAGTATGCGGAAATCAAGAGCGCTTATGATGA 1115
 QY 830 CGAGGAGAGCGCGGAGAGCTGAGAGCGCGCTGCTGCGGCAACCGCGCGCTGCGGCA 889
 Db 1116 ATCTGCTTGGCTTCACTTCCATATTAAGAGCTGAGAGAGAGAGAGAGAGAGAGAG 1175
 QY 890 CGCGGAGCTGCGGAGCTTCACTTCAATGATGATGATGATGATGATGATGATGATGAT 949
 Db 1176 TCTTCAGGCGCACTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1235
 QY 950 CCGCGGAGCGCGCTCTGCTGCGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAGG 1009
 Db 1236 TCGTCTTCTTCTCGGAG 1295
 QY 1010 CCACTGCTGCTCGCGGAG 1069

QY 1296 ACAGCTGCTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1355
 Db 1070 CGCCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1117
 QY 1356 GATGAGATTTTCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1415
 Db 1118 GAGAGAGCTGAGCGCTTCCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1177
 QY 1416 GATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1475
 Db 1178 CGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1237
 QY 1476 ACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1535
 Db 1238 GACAAAGTTCAAGTGG-----CGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGAG 1291
 QY 1536 GCTTCTTCTGCGAGATGCTTATCTCT 1562
 Db 1292 CATGCTGCTGAGATGAGCGCGCT 1318

RESULT 14
 ABT23282
 ID ABT23282 standard; DNA; 1585 BP.
 XX
 AC ABT23282;
 XX
 AC 01-MAY-2003 (first entry)
 DT
 XX
 DE Seed development embryo/endosperm size alteration DNA SEQ ID 92.
 XX
 KM Cytochrome P450; embryo; endosperm; size; seed development;
 XX antisense inhibition; co-suppression; transformed plant; gene; ds.
 XX
 OS Hordeum vulgare.
 XX
 PN WO200299063-A2.
 XX
 PD 12-DEC-2002.
 XX
 PF 04-JUN-2002; 2002WO-US17562.
 XX
 PR 05-JUN-2001; 2001US-295921P.
 XX
 PR 28-NOV-2001; 2001US-334317P.
 XX
 PA (DUPC) DU PONT DE NEMOURS & CO E. I.
 XX
 PI Cahoon RE, Heppard BP, Nagasawa N, Sakai H;
 XX
 DR WPI; 2003-229292/22.
 XX
 DR P-PSDB; AB26712.
 XX
 PT New isolated nucleic acid fragment for antisense inhibition or
 PT co-suppression of a cytochrome P450 polypeptide associated with
 PT controlling embryo/endosperm size during seed development in plants
 XX
 PS Claim 2; Page 122; 137pp; English.
 XX
 XX The invention relates to an isolated nucleic acid fragment comprising a
 CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or
 CC 17201 base pairs fully defined in the specification, or their similar and
 CC functionally equivalent fragment or subfragment. The nucleic acid encodes
 CC a cytochrome P450 polypeptide associated with controlling embryo/
 CC endosperm size during seed development. The nucleic acid fragment is
 CC useful in antisense inhibition or co-suppression of a cytochrome P450
 CC polypeptide associated with controlling embryo/endosperm size during
 CC development in a transformed plant. This polynucleotide sequence
 CC represents a gene relating to the alteration of embryo/endosperm size
 CC during seed development of the invention.
 XX
 XX Sequence 1585 BP; 228 A; 551 C; 523 G; 283 T; 0 other;
 Query Match 17.3%; Score 275.2; DB 25; Length 1585;

Best Local Similarity 54.7%; Pred. No. 2,4e-82;
Matches 706; Conservative 0; Mismatches 533; Indels 51; Gaps 6;

```

QY 306 CGAGCCCAAGAGGCTCANTGCTTTCACTTGAAGAGATCGCGTGAATGTCAGTCGCA 365
DB 289 CGGCGGCGCGGATGATGCTTCTCCGTCGCTCAACCCGCTGCTGCTGCTGCTGCTG 348
QY 366 TCCGACGTAAGGAAAGATTTCTGAATAGCCCGGTTTTTGTGATGCAACCGGTAAAGA 425
DB 349 GCAAGATACGCGCGCGGATGATGCTCAACCCGCGCTTCGCGACCGCGCGGTAAAGA 408
QY 426 ATCGGCTTACCTGATGATTTAAACAGACAAATGGTTTTCAC---CACACGATGTTA 482
DB 409 CCGGCGCGCGCACTCTCTTCCACCGCGCAATGGGTTTTTGCCTGCGCGCGACGCCCA 468
QY 483 CTGCGCAAGCTTCCGCGCTATCGCTTCGAACATCTCTTATGACAAACAAATGAGAG 542
DB 469 CTGCGCTGCGCTGCGCGCTCTCGCGCGCACTCTCGGCTCTCGCGCTGCGCGCT 528
QY 543 AGCCGAGACGCAACGACGATGAT---CTCAAGCCAGATGATGATTTCTTGAACAA 599
DB 529 CTCCGACCCCAACCGTCTCTATTTGGGCGGCGCAATGCTGGGCAAGTGCCTCATCAT 588
QY 600 GAGTATTAACGAACCTGTTTGTGATGATTTGATTAACGGCGTGGCTTAACACAT 659
DB 589 GGGCCCGCAACGCGAGTGCCTCTCGAGGTTCTGCAACGCGGCTCTCAACACAT 648
QY 660 GATGCTCTGATTTGCAACAGATGATGATGATGATGATGATGATGATGATGATGAT 713
DB 649 CATGCGCTGCTTTGCGCAAGCGCTACACACATTTCAACAGCAAGAGAGTGTGT 708
QY 714 AGCTGAATGCTGCAAGAGTTATGATTTGCTCGAACGTTGAATTTGATGATGATGAT 773
DB 709 GGAGAGATGCTTAAAGAGGATGACCTCTCGGCAAGTCTCACTGCGAGATCAGCT 768
QY 774 TCGTGGCTATGCGAATTTGATCTCAAGACCTCGGATGATGATGATGATGATGATGAT 833
DB 769 GCCATTCTCAATGCTGCTCATCTCAAGGCGCTGCGCGCGCGGTGCAACAGTATGCT 828
QY 834 AAGGTAACCGGTTTGTATCCGATTTATTCGAACACCGTATCAACACCG--- 887
DB 829 GCAAGTGAAGCGCTACGTCGTAACATCAACAGAGCAAGCGCGACAGCTGC 888
QY 888 -----TGATTTGCTGCTGATTTGCTGCGAGTTTCTCTCTCTCTCTCTCTCT 932
DB 889 ATGAGGCTTGCAGATGATGCTCTCGGCGACCTTCTGCTGCTCTCTCTCTCTCTCT 948
QY 933 TTGAGATTAATTAATCCGACCCGACATATGCGCGCTTTTGGAGATGATATTCAGAG 992
DB 949 AGAAGCAAGATGTCAGATGTCGACATGATGCGCGCTTTTGGAGATGATCTTTAGAG 1008
QY 993 AACAAGACAGTTGCGGCTTTATTCAGATGATCTCTGCTAGATGATGCTCTTATCAAA 1052
DB 1009 GACGGAACAGTGGCGATCTTGATGAGATTAAGCGAGATGATGATGATGATGATG 1068
QY 1053 TATGATCAACGCTGACAAACAGCTGATCAAGTATGCGGAAATCAAGAGCCCTAA 1112
DB 1069 GATTCAGTCAAGAGCCCGCGCGGAGATGACGCTGATGCGCGCGCGCGCGCGCTGAC 1128
QY 1113 TGAATCTGACTTGGCTTCACTTCATATCTTAAGCGCTGCTGTAAGAGATATGAGGCT 1172
DB 1129 GAGAGAGAGCTGTCAGAGCTCTCCATCAATCCAGTGCATGTAAGAGAGCGCTGAG 1188
QY 1173 TCAATCTCAAGGCGCACTTCTATCATGCGCGCTTTGGCAATTAACAGACATGCTTGA 1232
DB 1189 GCAACCGCGCGCGCTCTCTCTATGAGGCGCGCTGCGTCAACAGCGCGACCTTCG 1248
QY 1233 TGGTGTCTTGTTCGCGAGGACACAGATGATGATGATGATGATGATGATGATGATG 1292
DB 1249 CGGCACTGCTGCGCGCGCGCAACGAGATGATGATGATGATGATGATGATGATGATG 1308
QY 1293 TCCACAGTGTGGGTGATCTTTGAGTTTAACTGAGAGGTGCTGCGCAAAAGAG 1352

```

```

DB 1309 CGGCGCGCTGTCGCGCGGACCGGAGCTGTTCCGCGCGGACGCTGATG----- 1357
QY 1353 TGAGTGAATTTTCGTTCTTGGGTCGATTTGACACTTGAACCTTTGGGTCGGGTCG 1412
DB 1358 -GAGAGAGAGTGAAGCTGTCGAGCAGACCTTCCGCTTGGCTTGGCGCGCGCGCG 1416
QY 1413 TCGGATTTGCCCCGGAAGATCTTGTACTTACCGTTATGTTTGAACGCGCATGAT 1472
DB 1417 GCGGTGTGCCCCGGAAGATGCTGGCGCTGCGCACCGTCACTGCTCGCGACGCT 1476
QY 1473 GTTACATGATTTGAATGCGGACCGTCCGATGATTAACGCGCTGATCTTATCTGAGAACT 1532
DB 1477 GCTTCAACGCGTTCGATGAGTGGCTCTCTCG-----GAGCGTGAACCTGTCAAGAGCT 1530
QY 1533 GAGGCTTTCTTGCAGATGAGCTATCTCT 1562
DB 1531 CAAGATGTCACGTGAGATGCGCACGCGCT 1560

```

RESULT 15
ABT23238
ID ABR23238 standard; DNA; 1539 BP.

AC ABR23238;

DT 01-MAY-2003 (first entry)

DE Seed development embryo/endosperm size alteration DNA SEQ ID 18.

XX Cytochrome P450; embryo; endosperm; size; seed development;
XX antisense inhibition; co-suppression; transformed plant; gene; de.

XX Zea mays.

XX MO20029063-A2.

XX 12-DEC-2002.

XX 04-JUN-2002; 2002MO-US17562.

XX 05-JUN-2001; 2001US-295921P.

XX 28-NOV-2001; 2001US-334317P.

XX (DUPO) DU PONT DE NEMOURS & CO E. I.

PI Cahoon RE, Heppard BP, Nagasawa N, Sakai H;

XX WPI; 2003-229292/22.

DR P-PsDB; ABJ26682.

XX New isolated nucleic acid fragment for antisense inhibition or

PT co-suppression of a cytochrome P450 polypeptide associated with

PT controlling embryo/endosperm size during seed development in plants

XX Claim 38; Page 85; 137pp; English.

XX The invention relates to an isolated nucleic acid fragment comprising a

CC promoter consisting essentially of a sequence having 2000, 2022, 8300 or

CC 17201 base pairs fully defined in the specification, or their similar and

CC functionally equivalent fragment or subfragment. The nucleic acid encodes

CC a cytochrome P450 polypeptide associated with controlling embryo/

CC endosperm size during seed development. The nucleic acid fragment is

CC useful in antisense inhibition or co-suppression of a cytochrome P450

CC polypeptide associated with controlling embryo/endosperm size during seed

CC development in a transformed plant. This polynucleotide sequence

CC represents a gene relating to the alteration of embryo/endosperm size

CC during seed development of the invention.

XX Sequence 1539 BP; 274 A; 481 C; 485 G; 298 T; 1 other;

XX Query Match 17.2%; Score 274.4; DB 25; Length 1539;

XX Best Local Similarity 60.3%; Pred. No. 4.4e-82;

XX Matches 524; Conservative 0; Mismatches 312; Indels 33; Gaps 3;

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rng

Page 18

QY 709 GAGTTACGTAATAGTGTGAGAAAGATTGATTGCTCGGAAACGTTGAATTGACTGAT 768
DB 256 GAGCTGAGGCGACTGTGTGAGAGAGGAGACACTGCTGGGCGAGATCAACTGTGTCGAC 315
QY 769 CACCTTCTTGGCTATCGAGATTGATCTCAAGACTCCGGTCTAGATGTCCACTC 828
DB 316 CACCTCCCGGCGCTGCGTGTGCTGACCTGACAGACACAGAGGCGAGGTGACCGCTC 375
QY 829 GTACCAAGGTAAACCGGTTGTAATCCGATTATATCCGAACACCGTAAATCAACCGGT 888
DB 376 GTCCGCGCTGTGACCCGCTTGTGTCGGCGGCAATGTCAGACAGACACCGCGCCGGAACAC 435
QY 889 GATTGCTCTGT-----GATTGCTGAGGTTTGTCTCTCCCTCCATGAT 933
DB 436 CTCCGCTGTGCTCCGCTGCGCTGCTGCTGACTTCAACCGACTCTGCTCTGCTGCGGCC 495
QY 934 TCAGATAATATTCGACCCGGAATATGCGCCGTTCTTGGAGATGATATTCAGAGA 993
DB 496 GACGACAGGCTCACCGACGCTGACATGATCCGCTCTCTGGAAATGCTTCCGTGA 555
QY 994 ACAGACAGAGTTCGGTCTTAATCGAGTGAATCTGCTAGAGATGCTTCAATCCAGAT 1053
DB 556 ACTGACACGTCGCTGCTGATGAGTGGGCGCTGCGCAGGCTGCTGCTGACCTGAC 615
QY 1054 ATGCAATCAAGGTCAAAAGAGTGAATCAAGTGAAGGGAATCAAGAGCCCTAGAT 1113
DB 616 GTGACGCGCGGCTTCAAGAGAGTGAACCGCTGCTGCGGCCGACCGGCGCTCACC 675
QY 1114 GAATCTGAATTGGCTTCACTTCAATATTAACGGCTGTGTGAAGAGATTAAGGCTT 1173
DB 676 GAGTCCGACAGGGCTCACTGTGTACTGTGACGCGCTGATCAAGAGTGTCTCAGATG 735
QY 1174 CATCTCCAGGCGGCTTCTATCATGAGGCGGTTGGCCATTAAGACAGAGATGAT 1233
DB 736 CACCGCGGCGCGCTGTGTGTGGGGCGCTTGGCCAGCTGACGTCACGTCGAC 795
QY 1234 GGTGCTTGTTCGCGCAGGAGCACAGCAATGATGAACATGTGGCGGTATGCAATGAT 1293
DB 796 GGGCACTCATCCCGCGCGGACCAACCGGATGTGAACATGTGGCCATTACGACGAC 855
QY 1294 CCAACGCTGTGGTGTGATCTTGAAGTTAACTGAGAGTTGTGTGCAAAAGAGT 1353
DB 856 CCAAGACGTGTGGCGGAGCCGACGAGTTCACGCGGAGAGGTTGATG-----GGG 906
QY 1354 GAGTGAAGTTTTCGGTCTTGGGTGAGTTGAGACTTGCACCTTCCGGTCCGGTCT 1413
DB 907 TCCACCGAGTTCGCCATCATGAGGTGAGACCTCAGGCTGCGCCGTTCCGGGCGGGCCGG 966
QY 1414 CGAATTGCCCCGGGAAAGATCTTGGTTTACTACCGTTAGTTTGAACGCGCATGATG 1473
DB 967 CCGAGCTGCCCGGGAAGAGCTTCGCGCAATGAGCCACCGTGGCTTCTGCTGCGCAGCTG 1026
QY 1474 TTACATGAGTTGAATGAGGAGCCGTCGATGTTAAGGCGTTGACTTAATCTAGAACTG 1533
DB 1027 CTGCAAGAGTTGAGAGTGTCTCCCTC-----GCCGTGCACTGTGAGAGGTGCTC 1077
QY 1534 AGGCTTCTTGGAGATGAGTCTTCT 1562
DB 1078 AAGCTGTGTCGAGATGCGCTCCGCT 1106

Search completed: January 17, 2004, 04:23:05
Job time : 456 secs

RESULT 2
AV807688/c 436 bp mRNA linear EST 29-MAR-2002
LOCUS AV807688 Arabidopsis thaliana cDNA clone RAFL09-51-118 3',
DEFINITION MRSA sequence.
ACCESSION AV807688
VERSION AV807688.1 GI:19841673
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 436)
REFERENCE 1 (bases 1 to 436)
AUTHORS Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
Aizawa, T., Shibata, K., Shingawa, A., Muramatsu, M., Hayashizaki, Y.
and Shinozaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished
COMMENT Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyada, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msek@rtc.riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified Bluescript vector. Please visit our web
site (http://www.gen.riken.go.jp/e/plant/index_e.html) for further
details.

FEATURES
source
1..436
/location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL09-51-118"
/dev_stage="plants at various developmental stages from
germination to mature seeds"
/lab_host="DH10B"
/clone_11b="RAFL9"
/note="Site 1: BamHI; Site 2: SalI; subjected to
dehydration (1, 2, 5, 10, 24 hr) and cold (1, 2, 5, 10, 24
hr) treatments"

BASE COUNT 168 a 92 c 56 g 120 t
ORIGIN

Query Match 3.9%; Score 79; DB 9; Length 436;
Best Local Similarity 100.0%; Pred. No. 2,7e-28;
Matches 79; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1931 GATGGCTATCTCTTCGCTTAATGCGCGTAGGCGCAAGTAAAGAGCTCA 1990
DB 319 GATGGCTATCTCTTCGCTTAATGCGCGTAGGCGCAAGTAAAGAGCTCA 260
QY 1991 TATGAGAAATTAGAGATTTT 2009
DB 259 TATGAGAAATTAGAGATTTT 241

RESULT 3
LOCUS BZ380060 262 bp DNA linear GSS 26-NOV-2002
DEFINITION SALX_114536.29.00.x Arabidopsis thaliana TDNA insertion lines
survey sequence.
ACCESSION BZ380060
VERSION BZ380060.1 GI:25472537

KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 262)
REFERENCE Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadgilab
C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P.,
Zimmerman, J. and Ecker, J.R.
A sequence-indexed library of Insertion Mutations in the
Arabidopsis Genome
Unpublished
COMMENT Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of
TDNA. This sequence lies within an annotated intron of At2g46660.
Class: TDNA tagged.

FEATURES
source
1..262
/location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALX_114536.29.00.x"
/clone_11b="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines
each of which contains one or more TDNA insertion
elements. The resultant fragment for each line was
directly sequenced to determine the genomic sequence at
the site of insertion. Details of the protocols used can
be found at http://signal.salk.edu/cdna_protocols.html

BASE COUNT 64 a 56 c 62 g 75 t 5 others
ORIGIN

Query Match 2.5%; Score 50; DB 29; Length 262;
Best Local Similarity 100.0%; Pred. No. 1.3e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1495 ATGAATCTGACTTGGCTTCACTTCATATCTTAAGCGCTGTGTGAAGGA 1544
DB 160 ATGAATCTGACTTGGCTTCACTTCATATCTTAAGCGCTGTGTGAAGGA 209

RESULT 4
LOCUS BH451907/c 674 bp DNA linear GSS 12-DEC-2001
DEFINITION BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
survey sequence.
ACCESSION BH451907
VERSION BH451907.1 GI:17637618
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
1 (bases 1 to 674)
REFERENCE Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Frazer, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
COMMENT Other GSSs: BOGS044TF
Contact: Chris Town
TIGR Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES
source

1..674

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOG504"

/clone_1lb="BOGS"

/note="Vector: pHOSt, Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"

BASE COUNT
191 a 194 c 130 g 159 t

ORIGIN

Query Match 2.1%; Score 43; DB 28; Length 674;

Best Local Similarity 100.0%; Pred. No. 4.6e-10;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 CGCATGATCCACACGTGTGGTTCCTTGAATCCTTGAATTAACC 1711

DB 410 CGCATGATCCACACGTGTGGTTCCTTGAATCCTTGAATTAACC 368

RESULT 5

LOCUS

BH510149/c 745 bp DNA linear GSS 13-DEC-2001

DEFINITION BOGT049R BOGT Brassica oleracea genomic clone BOGT049, genomic

survey sequence.

ACCESSION

BH510149

VERSION

BH510149.1 GI:17718239

KEYWORDS

GSS.

SOURCE

Brassica oleracea

ORGANISM

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eustroids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 745)

Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of Brassica oleracea

Unpublished

Other GSS: BOGT049TF

Contact: Chris Town

TIGR 9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

Location/Qualifiers

1..745

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="TO1000DH3"

/db_xref="taxon:3712"

/clone="BOGT049"

/clone_1lb="BOGT"

/note="Vector: pHOSt, Site 1: BstXI, 2-3 kb sheared genomic DNA inserted into pHOSt using BstXI linkers"

BASE COUNT
219 a 186 c 137 g 203 t

ORIGIN

Query Match 2.1%; Score 43; DB 28; Length 745;

Best Local Similarity 100.0%; Pred. No. 4.6e-10;

Matches 43; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1669 CGCATGATCCACACGTGTGGTTCCTTGAATCCTTGAATTAACC 1711

DB 197 CGCATGATCCACACGTGTGGTTCCTTGAATCCTTGAATTAACC 155

RESULT 6

LOCUS

AL946977 82 bp DNA linear GSS 24-OCT-2002

DEFINITION Arabidopsis thaliana T-DNA flanking sequence GK-300D11-015600,

genomic survey sequence.

ACCESSION

AL946977

VERSION

AL946977.1 GI:24403599

KEYWORDS

GSS.

SOURCE

Arabidopsis thaliana

ORGANISM

Arabidopsis thaliana (thale cress)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eustroids II; Brassicales; Brassicaceae; Arabidopsis.

1

Strizhov, N., Li, Y., Rosso, M., Vlahov, P., Dekker, K., Saedler, H.

and Weisshaar, B.

A pipeline for automated high-throughput generation of FSTs

(flanking sequence tags) from Arabidopsis thaliana T-DNA

transformed lines

Unpublished

2

Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weisshaar, B.

A new Arabidopsis thaliana T-DNA mutagenesis population (GABI-Kat)

for flanking sequence tag based reverse genetics

Unpublished

3

(bases 1 to 82)

Li, Y., Strizhov, N., Rosso, M. and Weisshaar, B.

Direct Submission

Submitted (21-OCT-2002) Weisshaar B., Max-Planck-Institut fuer

Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany

This sequence is recovered from the left border of the T-DNA. It

indicates an insertion close to or within gene At2g46660. The

sequences are generated at the MPI for Plant Breeding Research in

the context of the GABI-Kat project. GABI-Kat is part of the German

Plant Genomics program designated 'GABI'. Information on line

availability can be found at:

http://www.mpiz-koeln.mpg.de/GABI-Kat/.

Location/Qualifiers

1..82

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia 0"

/db_xref="taxon:3702"

/clone="GK-300D11-015600"

/note="PCR was performed on DNA from Arabidopsis thaliana

plants (T1) which were transformed with the T-DNA from

vector pACT61. The lines contain one or more T-DNA

insertions. The DNA fragment(s) resulting from the PCR

were directly sequenced to determine the genomic sequence

flanking the insertion. Sequences displaying significant

similarity to the A. thaliana nuclear genome sequence were

processed for submission. T-DNA derived sequences were

removed"

BASE COUNT
17 a 13 c 12 g 19 t 21 others

ORIGIN

Query Match 1.9%; Score 38; DB 29; Length 82;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1379 AGACACAGTTCGGTCTTAATGAGTGATCTTGGCTA 1416

DB 45 AGACACAGTTCGGTCTTAATGAGTGATCTTGGCTA 82

RESULT 7

LOCUS

B2428529 809 bp DNA linear GSS 13-DEC-2002

DEFINITION BONHW74TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONHW74,

genomic survey sequence.

ACCESSION

B2428529

```

VERSION B2428529.1 GI:26671510
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 809)
Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BONHW74TF
COMMENT Contact: Chris Town

FEATURES
    source
        1. 809
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="T01000DH3"
            /db_xref="taxon:3712"
            /clone="BONHW74"
            /note="Vector: PHOS1; Site 1: BstXI, 1.6-2 kb sheared
            total DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 241 a 216 c 146 g 206 t

ORIGIN
Query Match 1.8%; Score 37; DB 29; Length 809;
Best Local Similarity 100.0%; Pred.No. 5e-07;
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1627 TTCCGCGAGGAGCAGCAGATGATGATGAGC 1663
DB 473 TTCCGCGAGGAGCAGCAGATGATGATGAGC 437

RESULT 8
B2428523 712 bp DNA linear GSS 13-DEC-2002
LOCUS BONHW74TF BO.1.6.2_KB tot Brassica oleracea genomic clone BONHW74,
DEFINITION genomic survey sequence.
ACCESSION B2428523
VERSION B2428523.1 GI:26671493
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 712)
Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other_GSSs: BONHW74TR
COMMENT Contact: Chris Town

FEATURES
    source
        1. 712
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="T01000DH3"
            /db_xref="taxon:3712"
            /clone="BONHW74TR"
            /note="Vector: PHOS1; Site 1: BstXI, 1.6-2 kb sheared
            total DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 241 a 182 c 155 g 226 t

ORIGIN
Query Match 1.6%; Score 32; DB 28; Length 804;
Best Local Similarity 100.0%; Pred.No. 0.00017;
Matches 32; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 995 AAGTAAACCGGTTGATCCCGATATATC 1026
DB 545 AAGTAAACCGGTTGATCCCGATATATC 514

Search completed: January 16, 2004, 19:16:55
Job time : 4299 sec

FEATURES
    source
        1. 804
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="T01000DH3"
            /db_xref="taxon:3712"
            /clone="BOHR146"
            /note="Vector: PHOS1; Site 1: BstXI, 2-3 kb sheared
            genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 241 a 182 c 155 g 226 t

ORIGIN
Query Match 1.7%; Score 35; DB 29; Length 712;
Best Local Similarity 100.0%; Pred.No. 5.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 AAACGGCGTCGCTTACACATGATGCTGT 831
DB 387 AAACGGCGTCGCTTACACATGATGCTGT 421

RESULT 9
B2431683/c 804 bp DNA linear GSS 12-DEC-2001
LOCUS BOHR146TR BOHR Brassica oleracea genomic clone BOHR146, genomic
DEFINITION survey sequence.
ACCESSION B2431683
VERSION B2431683.1 GI:17617404
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 804)
Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Contact: Chris Town

FEATURES
    source
        1. 804
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="T01000DH3"
            /db_xref="taxon:3712"
            /clone="BOHR146"
            /note="Vector: PHOS1; Site 1: BstXI, 2-3 kb sheared
            genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 241 a 182 c 155 g 226 t

ORIGIN
Query Match 1.7%; Score 35; DB 29; Length 712;
Best Local Similarity 100.0%; Pred.No. 5.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 AAACGGCGTCGCTTACACATGATGCTGT 831
DB 387 AAACGGCGTCGCTTACACATGATGCTGT 421

RESULT 9
B2431683/c 804 bp DNA linear GSS 12-DEC-2001
LOCUS BOHR146TR BOHR Brassica oleracea genomic clone BOHR146, genomic
DEFINITION survey sequence.
ACCESSION B2431683
VERSION B2431683.1 GI:17617404
KEYWORDS GSS.
SOURCE Brassica oleracea
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 804)
Town,C.D., Van Aken,S., Uteback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Contact: Chris Town

FEATURES
    source
        1. 804
            /organism="Brassica oleracea"
            /mol_type="genomic DNA"
            /strain="T01000DH3"
            /db_xref="taxon:3712"
            /clone="BOHR146"
            /note="Vector: PHOS1; Site 1: BstXI, 2-3 kb sheared
            genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 241 a 182 c 155 g 226 t

ORIGIN
Query Match 1.7%; Score 35; DB 29; Length 712;
Best Local Similarity 100.0%; Pred.No. 5.2e-06;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 797 AAACGGCGTCGCTTACACATGATGCTGT 831
DB 387 AAACGGCGTCGCTTACACATGATGCTGT 421

```


Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rn1

Page 1

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2004, 01:32:45 : Search time 104 Seconds
(without alignments)
6760.801 Million cell updates/sec

Title: US-10-022-025a-2
Perfect score: 1593
Sequence: 1 atgctcgaactcgaag.....tgcgcctagcgcagttaa 1593
Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues
Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA: *
1: /cgm2_6/ptodata/1/ina/5A COMB.seq: *
2: /cgm2_6/ptodata/1/ina/5B COMB.seq: *
3: /cgm2_6/ptodata/1/ina/6A COMB.seq: *
4: /cgm2_6/ptodata/1/ina/6B COMB.seq: *
5: /cgm2_6/ptodata/1/ina/PTUS COMB.seq: *
6: /cgm2_6/ptodata/1/ina/backfile1.seq: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	651.6	40.9	1611	3 US-08-948-564-7	Sequence 7, Appl1
2	355	22.3	1799	2 US-08-560-398-3	Sequence 3, Appl1
3	144.2	9.1	1704	3 US-08-948-564-17	Sequence 17, Appl1
4	129.6	8.1	1708	3 US-08-991-677-1	Sequence 1, Appl1
5	98.2	6.2	275	4 US-09-313-294A-2591	Sequence 2591, Ap
6	93.6	5.9	1788	3 US-08-948-564-9	Sequence 9, Appl1
7	93.6	5.9	1893	1 US-08-532-065B-1	Sequence 1, Appl1
8	88.6	5.6	1761	3 US-09-033-055A-3	Sequence 3, Appl1
9	86.8	5.4	2174	3 US-08-606-505B-63	Sequence 63, Appl1
10	86.8	5.4	2174	3 US-09-616-990-63	Sequence 63, Appl1
11	86.2	5.4	1927	3 US-08-606-505B-64	Sequence 64, Appl1
12	86.2	5.4	1927	3 US-09-616-990-64	Sequence 64, Appl1
13	79.6	5.0	1657	3 US-08-948-564-11	Sequence 11, Appl1
14	78.6	4.9	1835	4 US-09-564-808-5	Sequence 5, Appl1
15	78.6	4.9	1838	4 US-09-091-432-1	Sequence 1, Appl1
16	78.6	4.9	1838	4 US-09-387-663-1	Sequence 1, Appl1
17	77.2	4.8	5156	2 US-09-091-432-3	Sequence 3, Appl1
18	77.2	4.8	5156	2 US-09-387-663-3	Sequence 3, Appl1
19	75.6	4.7	384	4 US-09-615-192A-232	Sequence 232, App
20	75.4	4.7	1722	3 US-09-033-055A-1	Sequence 1, Appl1
21	75.4	4.7	1755	1 US-07-912-900-29	Sequence 29, Appl1
22	75.2	4.7	1755	1 US-08-285-309-29	Sequence 29, Appl1
23	75.2	4.7	1755	1 US-08-502-046-29	Sequence 29, Appl1
24	75.2	4.7	1812	1 US-08-313-075A-37	Sequence 37, Appl1
25	75.2	4.7	1824	3 US-08-606-505B-1	Sequence 1, Appl1
26	75.2	4.7	1824	3 US-09-616-990-1	Sequence 1, Appl1
27	74.4	4.7	1620	4 US-09-627-216A-11	Sequence 11, Appl1

28	74.4	4.7	1620	4 US-09-765-873A-11	Sequence 11, Appl1
29	73.8	4.6	1884	4 US-09-564-808-3	Sequence 3, Appl1
30	72.8	4.6	543	4 US-09-615-192A-230	Sequence 230, App
31	72	4.5	1737	4 US-09-126-420A-1	Sequence 1, Appl1
32	70.6	4.4	1880	4 US-09-564-808-1	Sequence 1, Appl1
33	70.6	4.4	1634	4 US-09-126-420A-2	Sequence 2, Appl1
34	67	4.2	1665	1 US-08-313-075A-29	Sequence 29, Appl1
35	66.6	4.2	1824	3 US-08-948-564-13	Sequence 13, Appl1
36	65.2	4.1	1866	4 US-09-615-192A-103	Sequence 103, App
37	65.2	4.1	2013	4 US-09-615-192A-104	Sequence 404, App
38	64.4	4.0	1506	3 US-09-158-767-7	Sequence 7, Appl1
39	64.4	4.0	1506	3 US-09-158-767-8	Sequence 8, Appl1
40	64.4	4.0	1506	3 US-09-158-767-9	Sequence 9, Appl1
41	64.4	4.0	2261	3 US-09-351-229-3	Sequence 1, Appl1
42	64.2	4.0	1806	4 US-09-351-229-3	Sequence 2, Appl1
43	64.2	4.0	1812	1 US-07-912-900-28	Sequence 28, Appl1
44	64.2	4.0	1812	1 US-08-285-309-28	Sequence 28, Appl1
45	64.2	4.0	1812	2 US-08-502-046-28	Sequence 28, Appl1

ALIGNMENTS

RESULT 1
US-08-948-564-7
Sequence 7, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Simlinszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Cordin, Frederick T.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constrictants and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSEE: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948, 564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 1611 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 20..1588
US-08-948-564-7
Query Match 40.9%; Score 651.6; DB 3; Length 1611;
Best Local Similarity 66.5%; Pred. No. 2.9e-218;
Matches 1025; Conservative 0; Mismatches 484; Indels 33; Gaps 5;

QY 61 AGCCAAACCACTTGGCTTCTCCCTCCGCTCAATCATCTGGCTGCGCATCT 120
 Db 74 ACCCAAGAAAACCTTGATGGGTCTTTTATCATAGGCTCACTCTGTATACATGACT 133
 QY 121 CTCTTCTTATGACCTTATCCCGGTGACCTGCTGGGGAATACCTCTTCCGCGGTTA 180
 Db 134 TTTATTAATCTGTACACCTCGGTGTCTGCTGGGCAATCATACAC-----TA 186
 QY 181 ATATCCGGTATACAAACCGAAACGTTATTCGCGTCCAAAGGCTTCCCTTGGTT 240
 Db 187 CTCTCCCCCTTTCAA-----TCAATCCGGTCCCAAGGCTTCCCTTATT 235
 QY 241 GGAAGCATGCTACTATGCTCAAGCACTTACTGACCAAGATCGCTGATGACGCTGAG 300
 Db 236 GGAACCATGGGCTTATGACTTCC---CTGGCCCATCACTGATCGACGCGCGCGCC 292
 QY 301 AAATCGAGCCCAAGAGGCTCATGGCTTTTCACTTGAAGAGACTCGCGTATGCTACG 360
 Db 293 AATGCAAGAGCCCAAGGCTCATGGCTTTTATGCTCGCGCAACACGCTGATGCTACG 352
 QY 361 TGCATCCCGAGTGAAGCAAGATTTGTAATAGCCCGGTTTTTGTGATGACCGGTT 420
 Db 353 TGCCACCCCGAGTGAAGCAAGATTTGTAATAGCCCGGTTTTTGTGATGACCGGTT 412
 QY 421 AAAGAAATCGGCTTACTCACTGATGTTTAAAGAGCAATGGTTTTGACACACAGGTT 480
 Db 413 AAAGAAATCGGCTTACTCACTGATGTTTAAAGAGCAATGGTTTTGACACAGGTT 472
 QY 481 TAATGCGCAAGCGTTCCGCGTATCGCTTGAACCATCTCTTAAAGTAAACAAATGAG 540
 Db 473 TAATGCGCAAGCGTTCCGCGTATCGCTTGAACCATCTCTTAAAGTAAACAAATGAG 532
 QY 541 AGAGCCGAGCCCAAGAGTGAATGCTCAAGCCAGATGTTGTTCTTGAAGAAACAG 600
 Db 533 GCTTGTAGCTCCAAAGCTCTCAATGCGCGCCCAAGATGTTTCAATCT---AAATAC 588
 QY 601 AGTATTAACGAACCTGTTTTTGTGTAAGTTGCTTAAACGCGCTGCTTAAACAGATG 660
 Db 599 AAGCCCAACCGAGCTTATGCTGTGCGCAAGTGTGAAGAGGCTTCCCTCATGATG 649
 QY 661 ATGTGCTGTATTTGGCAACAGATGATGAGCTGAAAAACCA-----TGTGAG 711
 Db 650 ATGTGCTGTATTTGGCAACAGATGATGAGCTGAAAAACCA-----TGTGAG 709
 QY 712 TTAAGTAAATGATGTAAGAGATGATGTTGCTCGAAGCTTGAATGATGATGATG 771
 Db 710 CTGGAATATGATGTAAGAGATGATGTTGCTCGAAGCTTGAATGATGATGATG 769
 QY 772 CTTCCTTGGCTATCGAGTTTATCTCTCAAAAGCTCGGTCTAATGTTTCCACACTGTA 831
 Db 770 CTTCCTTGGCTATCGAGTTTATCTCTCAAAAGCTCGGTCTAATGTTTCCACACTGTA 829
 QY 832 CCAAGAGTAAACCGGTTGATCGGATTAATATCGAACAACCGTAAATCAACCGGTTAT 891
 Db 830 CCAAGAGTAAACCGGTTGATCGGATTAATATCGAACAACCGGTTAAATCAACCGGTT 889
 QY 892 TTGCTGTGATTTGCTGAGCTTTTGTCTCTCTCCCTCAATGTTGATGATGATG 951
 Db 890 ACCAATCGGATTTGCTGAGCTTTTGTCTCTCTCCCTCAATGTTGATGATGATG 949
 QY 952 CCGGACATATCGCGCTTTCTTGGGAGATGATTAATCAAGAGACAGACAGTGGCTG 1011
 Db 950 TCCGACATATCGCGCTTTCTTGGGAGATGATTAATCAAGAGACAGACAGTGGCTG 1009
 QY 1012 TTAATGATGATCTCTGCTGAGATGATGCTTCAATCGATGATGATGATGATG 1071
 Db 1010 TTGATGATGATCTCTGCTGAGATGATGCTTCAATCGATGATGATGATGATG 1069
 QY 1072 AACGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1131
 Db 1070 GAGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1129
 QY 1132 CTTCATATCTAAAGGCTGTGTGTAAGAAAGATTTGAGGCTTCAATCTTCCAGGCCACTT 1191

Db 1130 ATGAGTATCTAACCGCGGTGTGTAAGAGGTTGCTGCGCTGACCCCGCGGCCACTT 1189
 QY 1192 CTATCATGGGCGCGTTTGGCCATAACAGACAGATGTTGATGTTGCTGTTCCGCA 1251
 Db 1190 CTATCATGGGCGCGTTTGGCCATAACAGACAGATGTTGATGTTGCTGTTCCGCA 1249
 QY 1252 GGAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1311
 Db 1250 GGAACCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1309
 QY 1312 CTTTGAAGTTTAACTGAGAGTTCGTCGCAAAAGAGTGAAGTGAAGTTTGGTT 1371
 Db 1310 CCACTGAAATTTATGCTCCGAGAGTTTGTCACTGGGGGTGAGATGCCGAATTTT 1369
 QY 1372 CTGGGTGGAATTTGAGACTTGCACCTTTGCGGTGCGTCTGCGATTTGCCCGGAG 1431
 Db 1370 CTGGGTGGAATTTGAGACTTGCACCTTTGCGGTGCGTCTGCGATTTGCCCGGAG 1429
 QY 1432 AATCTGTTTACTACGCTTATGTTTGAAGGAGATGATGATGATGATGATGATG 1491
 Db 1430 ACTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1489
 QY 1492 GGAACCTCCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1551
 Db 1490 GTACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1549
 QY 1552 GCTAATCTCTTCCGCTAATTTGGCGGTGAGCGGATGATGATGATGATGATG 1593
 Db 1550 GCTAATCTCTTCCGCTAATTTGGCGGTGAGCGGATGATGATGATGATGATG 1591

RESULT 2
 US-08-560-398-3
 ; Sequence 3, Application US/08560398
 ; Patent No. 5907082
 ; GENERAL INFORMATION:
 ; APPLICANT: O'Neill, Sharmen
 ; APPLICANT: Nadeau, Jeanette
 ; TITLE OF INVENTION: Ovine-Specific Gene Expression
 ; NUMBER OF SEQUENCES: 13
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Townsend and Townsend and Crew LLP
 ; STREET: Two Embarcadero Center, Eighth Floor
 ; CITY: San Francisco
 ; STATE: California
 ; COUNTRY: USA
 ; ZIP: 94111-3834
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patentin Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/560,398
 ; FILING DATE: 17-NOV-1995
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Baetian, Kevin L.
 ; REGISTRATION NUMBER: 34,774
 ; REFERENCE/DOCKET NUMBER: 023070-063300US
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (415) 576-0200
 ; TELEFAX: (415) 576-0300
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1799 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; FEATURE: DNA (genomic)
 ; NAME/KEY: unsure

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rnt

Page 3

LOCATION: 19..23
FEATURE:
NAME/KEY: CDS
LOCATION: 291..1571
OTHER INFORMATION: /note= "clone 040 ovule-specific gene
OTHER INFORMATION: encoding a cytochrome P450 monooxygenase
OTHER INFORMATION: from pollen tubes of Phalaenopsis"
US-08-560-398-3

Query Match 22.3%; Score 355; DB 2; Length 1799;
Best Local Similarity 58.0%; Pred. No. 6,4e-114;
Matches 744; Conservative 0; Mismatches 500; Indels 39; Gaps 5;

QY 304 TTGAGAGCCAAAGAGCTCATGCTTTCACTTACCTTACGAGAGAGACTGCGGTGATGCTCAAGTGC 363
DB 273 TTGAGAGCCCTCCGCTCATGCTTTCTGCTGCGCTTCACTGCTTCACTGCTTCAAGC 332
QY 364 AATCCGAGCTAGCGAAAGATTCGAATPAGCCGCTTTTGTGATGACCGGTAA 423
DB 333 CACCCGAAACCGCAAAAGATTTCTTCAAGCCGCTTGCCTGATGCGCCATTAA 392
QY 424 GAATGCGCTTACTCATGATGTTTAAAGAGCAATGCTTTTGCACACAGGTGTTAC 483
DB 393 GAATGAGCATACGAATCTGTTTAAACGGCTATGGGTTTGGCCCATTTGGGAGTAC 452
QY 484 TGGCGAAGCTTCCGCTATGCTTGAACCATCTCTTAGTACAAAACAAATCAGAGA 543
DB 453 TGAAGAAACCTGAGAAAGATTTGTCACATATCTTTCACTGCGCGAGTTTCATCG 512
QY 544 GCCGAGCGCAACGAGATGATCTCAAGCCAGATGATGATTTCTGAAAAACAGA-- 601
DB 513 TTGAGAGAGCAACGAGTGAATTTGGCGAAGAAATGCTGGGATATGAAAAGATGATG 572
QY 602 -GTAATACGAAACCTGCTTTTGTCTGAGTCTTAAAGCGCGTCTTAAACAGATG 660
DB 573 GAGAGAAATGAGATTTGTAAGAGAGAGAAATGTTCACTACGCGGTCTTTGAAATACATC 632
QY 661 ATGAGCTGCTATTCGGAAGAGTATGAGCTTGAATAAAACCATGCTTGAAGTAACTGAA 720
DB 633 ATGTTGACTTTTGGGAAAAAGTTGATTTTTCGAAAGATAGAGGGTTGAGCTTGAAG 692
QY 721 ATGCTC--GAAAGAGTATGATTTGCTCGGAAAGCTTGAATGACTGATCACTTCTCT 777
DB 693 TTGATCTTAAGGAAGATATGATTTCTTGGATCTTCACTGCGGCTGATCTTTGGCT 752
QY 778 TGGCTATCGAGTTTGAATCTTCAAGAGCTCGGCTATGATTTCCACATCGTACCAAG 837
DB 753 CTTTGGAGATGATTTGAAAGTTCGAAAGTGTGAGAGAAAGATGCAAACTTGTGATAG 812
QY 838 GTAACCGCTTTGATCCGAGATTAATCCGAAACCGTATCAAAAC----- 885
DB 813 GTCAATGATTTTGAAGAAAGATCATAGACAGATTAAGAGAGCAACGCGTACGAG 872
QY 886 -----GATATTTGCTCGTATTTGCTGACGTTTGTCTTCCCTCCATGCTTCAAT 939
DB 873 ATTGATAGAGGTGAAGGTGAAGATTTGATGTGCTTTGTTGGAGGAGAAAGAT 932
QY 940 AATATTCGACCCGCAATATGCGCTTCTTGGAGATGATTAATCAGAGAAACGAG 999
DB 933 AGACTCTCAGATATGATATGATGCTGAGTCTTTGGGAAATGATCTTTAAGGAACTGAT 992
QY 1000 ACACTTGGGCTTTAATCAGATGATCTTCTGATGATGCTTCTTCACTCAATATGCAA 1059
DB 993 ACTGTTGCCATCCATTTGAGATGACGTTGCTAGAAATGATTTCTTCACTGATATTCAA 1052
QY 1060 TCAACGCTAACAACGAGCTGATCAAGTATCGGAAATCAAGAGCCATAGATCT 1119
DB 1053 TCGAAGGCAAGAGTGAATGATTTGCTGCTTCACTTCAAGCCAGATTTGATCT 1112
QY 1120 GACTTGGCTCACTTCATATCTAACGCTGTGTGTAAGAAAGATTTGAGGCTTCACTCT 1179
DB 1113 GATATCCAAAGACTTCTTATCTCAATCTATAGTAAAGAAACCTTCAATGACATCTT 1172

QY 1180 CCAGGCCCACTTCTATCATGAGGCCGTTTGGCCATAAGACAGATCTGATGCTGT 1239
DB 1173 CCTGGGCTCTATGATGATGAGGCTGCGCTTACATCATACGTTCTGTTGATGCTAC 1232
QY 1240 CTGTCCGAGAGGAGCAACAGCAATGTAATGATGAGGCGCTATGATGATCAAC 1299
DB 1233 ATGATTTCTGCTGAGAGACTGCAATGATGATGATGATGATGATGATGATGATG 1292
QY 1300 GTGTGGTATATCTTTGAGTTTAACTGAGAGTGTGTGCAAAAGAGTGAAGCTG 1359
DB 1293 AACTGGCTGAGCTTAACTAACTGATGATGATGATGATGATGATGATGATGATGATG 1344
QY 1360 GAGTTTCCGTTCTGAGTGGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTT 1419
DB 1345 ---TCAATATCTTGTGCTGATTTAAGTTGAGATTTGAGATTTGAGATTTGAGATTT 1400
QY 1420 TGCCCGGAGAAATCTTGTGTTTACTACCGCTTATGTTTGAACGCGAGATGATTAAT 1479
DB 1401 TGCCCGGAGAAACGATGAGATTTGAGATTTGAGATTTGAGATTTGAGATTTGAGATTT 1460
QY 1480 GAGTTTGAATGAGAGAGCTGCAATGATGATGATGATGATGATGATGATGATGATG 1539
DB 1461 AGCTTCAATATGCTTCTTC--GAGAAATGATGATGATGATGATGATGATGATGATG 1517
QY 1540 TCTTGCAGATGAGCTTAATCTCT 1562
DB 1518 TCTCTGAGATGAAGATCTTT 1540

RESULT 3
US-08-948-564-17
Sequence 17, Application US/08948564
Patent No. 6121512
GENERAL INFORMATION:
APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Corbin, Frederick T.
TITLE OF INVENTION: No. 6121512a1 Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Virginia C. Bennett
STREET: PO Box 37428
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 1704 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 38..1564

US-08-948-564-17

Query Match 9.1%; Score 144.2; DB 3; Length 1704;
 Best Local Similarity 53.5%; Pred. No. 8.3e-40;
 Matches 355; Conservative 0; Mismatches 293; Indels 15; Gaps 2;

QY 866 CCGAACCCGTAATCAACCGGTGATTTGCTCGTATTTTCGTGACGTTTGTCTCC 925
 DB 813 CTGAGACACGCAAAATCTGTGTGTCAGAACATTTTGTGATGCTCTCCAT 872
 QY 926 TCCATGCTTCAATTAATTAATCCGACCGGACATATCGCGTCTTTGGAGATGAT 985
 DB 873 TCGAAGCAAAATTAAGCTTATAGTAAGACATCATTTGCTCTTTGGAGATGAT 932
 QY 986 TCAAGAACACACACGCTTGGCTTTAATCGATGATCTTCGTAGATGTCCTTC 1045
 DB 933 CAGCAGGATGACACACATGCAATTTCAATGATGGCCATGGCTGATGATAGAA 992
 QY 1046 ATCCAGATATGCAATCAACGGTACAAAACGAGTGTGATCAAGTATCGGGAATCAAG 1105
 DB 993 ACCCAAGGTCAACAAAGATCTCAAGAGCTAGACAGGTATTTGGCTTGAAGGG 1052
 QY 1106 CCTAGATGAATCTGACTTGGCTTCACTTCATATCTTAACGGCTGTGAAAGATAT 1165
 DB 1053 TGATGACTGAAGCAGACTTCTCAATCTCCCTTACTTCAATGTGACCAAGAACAA 1112
 QY 1166 TGAAGCTTCACTCCGACGCGCCCTTCTATGATGGGCCCTTTGGCCATAACAGACGA 1225
 DB 1113 TGAAGCTTCACTCCGACGCGCCCTTCTATGATGGGCCCTTTGGCCATAACAGACGA 1169
 QY 1226 TCGTATGATGCTGCTTGTTCGCGACGGAACACAGCATGTGAACATGTGGCCGTAT 1285
 DB 1170 AAGTTGAGGCTATGACATTCCTCAAGAGGCTCAATGTGATGATGATGTGGCGGTG 1229
 QY 1286 CGCATATCCACAGCTGTGGGTGATCTTTGGATTTAACTGAGAGGTTCTGTGCA 1345
 DB 1230 CCGGACCCGCGCTGTGAGAGATCCATGAGATTCGACCGGAAGGTTCTTTAAG 1289
 QY 1346 AAGAGGTAGAGTGAAGTTTCCGTTCTTGGGTGCGATTTGACATTCGCGGT 1405
 DB 1290 AGGATGATGACATGAAG-----GGCATGACTTAAAGCTTACATTCGCGGT 1337
 QY 1406 CGGCTGTGCGATTTGCCCCCGGGAAGATCTTGTGTTTACTACCGTATGTTTGACGG 1465
 DB 1338 CGGCTGACGAGTATGCCCCGGGTGCCCACTTGTATCACTTGGACAGATCCATGTTGG 1397
 QY 1466 CGATGATGTTACATGATTTGAATGGGACCGTCGATGTATACGCGCTTGAATTTCTG 1525
 DB 1398 GGCACCTCTTGACCATTTCTGTGTGACCCCACTGAGAGAAATGAAGCTGAGGAATTG 1457
 QY 1526 AGA 1528
 DB 1458 ACA 1460

RESULT 4

US-08-991-677-1
 ; Sequence 1, Application US/08991677A
 ; Patent No. 6252135
 ; GENERAL INFORMATION:
 ; APPLICANT: Chiang, Vincent L
 ; APPLICANT: Carraway, Daniel T
 ; APPLICANT: Smeltzer, Richard H
 ; TITLE OF INVENTION: Production of Syringyl Lignin in Gymnosperms
 ; FILE REFERENCE: 50617
 ; CURRENT APPLICATION NUMBER: US/08/991,677A
 ; EARLIER FILING DATE: 1997-12-16
 ; EARLIER APPLICATION NUMBER: US 60/033,381
 ; NUMBER OF SEQ ID NOS: 11
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1708

TYPE: DNA
 ORGANISM: Liquidambar styraciflua
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (48)..(1571)
 US-08-991-677-1

Query Match 8.1%; Score 129.6; DB 3; Length 1708;
 Best Local Similarity 52.8%; Pred. No. 1.1e-34;
 Matches 335; Conservative 0; Mismatches 284; Indels 15; Gaps 2;

QY 873 CCGTATCAAAACCGGTGATTTGCTCGTATTTGCTGACGTTTGTCTTCCCTCATG 932
 DB 827 CCGTAAAAAGATGTGTGAGCCCAACATATTCGTGATCATTTGCTCACCTTACAGA 886
 QY 933 TTCAGTAAATTAATCCGACCGGACATATATGCGCTTTTGGGAGATGATTCAGAG 992
 DB 887 GAATATGACTTAACTGAGAGACATATATTTGGGCTCTTTGGGATATGATCATGACG 946
 QY 993 AACAGACAGTGTGCTCTTATGAGTGTCTCTGATGATGATGATGATGATGATGATGAT 1052
 DB 947 CATGACACACCGCAATCTGTGAAATGGCCATGCGCGATTAATTAAGAACCAAG 1006
 QY 1053 TATGCAATCAACGTAACAAACGAGCTGATCAATGATCGGGAATCAAGCCCTGA 1112
 DB 1007 GGTGCAACAAAAGCTCAAGAGAGCTAGACATGATGATGATGATGATGATGATGAT 1066
 QY 1113 TGAATCTGACTTGGCTTCACTTCATCTTAACGCTGTGTAAGATGATGAGCT 1172
 DB 1067 CGATTTGACTTCTGAGCTCTTATCTTACATGATGATGATGATGATGATGATGATGAT 1126
 QY 1173 TATCTCTGAGCCCACTTATCATGAGCCCGTTTGGCCATACAGACAGATGTTGA 1232
 DB 1127 GACCCCTCA---ACACCATATGCTCCCTCATGCGCGATGCAACGTCMAATTTG 1183
 QY 1233 TGGTGTCTTGTTCGCGACGGAACACAGCATGTGTAACATGTGGCGTATGCTGATGA 1292
 DB 1184 TGGCTACGACATCTCTTAAGGATCAAAATTTATGATTAATGCTGGCGGTGCTGTGA 1243
 QY 1293 TCCACAGTGTGGGTGATCTTGTGATTTAACTGAGATTTGCTGTGGCAAGAGG 1352
 DB 1244 TCCAGAGATGTGGCGTGAACCACTAGATTTGACCGGAACGTTCTCTGAAGAGATGT 1303
 QY 1353 TGAAGTGAATTTGCGTCTTGTGCGATTTGAGACTTGACCTTCGCGGTGGGTGCG 1412
 DB 1304 CGACATGAAA-----GTCACATTAATAGCTACTGCGGTGTTGGTCAGGAG 1351
 QY 1413 TCGATTTGCCCCGGAAGATCTTGTGTTTAACTACCGTTATGTTTGGACGGCATGAT 1472
 DB 1352 GCGTGTTCGCCCGGTGACCACTTGGCATTAATTTGGTCATCATGATGATGATGATGAT 1411
 QY 1473 GTTACATGATTTGAATGGGACCGTCCATGAT 1506
 DB 1412 ATTGACCAATTTCTATTTGAGCCCTCTTAAGGT 1445

RESULT 5

US-09-313-294A-2591
 ; Sequence 2591, Application US/09313294A
 ; Patent No. 6476212
 ; GENERAL INFORMATION:
 ; APPLICANT: Lalundi, Raghunath V.
 ; APPLICANT: Ito, Laura Y.
 ; APPLICANT: Sherman, Bradley K.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
 ; FILE REFERENCE: PI-0017 US
 ; CURRENT APPLICATION NUMBER: US/09/313,294A
 ; EARLIER FILING DATE: 1999-05-14
 ; NUMBER OF SEQ ID NOS: 7600
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 2591
 ; LENGTH: 275
 ; TYPE: DNA

```

; ORGANISM: Zea mays
;
; FEATURE:
; NAME/key: misc feature
; OTHER INFORMATION: Incyte ID No. 6476212 700552946H1
;
US-09-313-294A-2591

```

Query Match	6.2%;	Score 98.2;	DB 4;	Length 275;
Best Local Similarity	64.7%;	Pred. No. 3.4e-24;		
Matches 178; Conservative	0;	Mismatches 93;	Indels 4;	Gaps 2

QY 1189 GGGCCCGTTGGCCCTTAACAGACAGATCGTTGATGTCGCTTGTTCGGCAGGACCA 1258

Db 5 GGGCCCGCTCCCACTCAAGCGTGCAGGTAGGGGGGTACTTGTGCGCGGGACCA 64

QY 1259 CAGCATGGTGAACATGTGGCCGTTATGGCATATCCACAGTGTGGGTATCTTTGG 1318

Db 65 CCGCATGGTGAACATGTGGCCCAATACCAATCCAGACCTGTGGCCTGAGCCAAATG 124

QY 1319 AGTTAAACCTGAGAGGTTCGTGGCAAAAGAGTGAGTGAGTTTCGGTTCTTGAGT 1378

Db 125 AGTTAGGCCCCAGAGGTTTCATAGGGCCCTGGCGGCCGA---GGAAGTCCGAAATATGGGTT 181

QY 1379 CGGATTTGAGACTTGCACCTTTGGGGTGGGGTCTGTGGATTTTGCCCCGGGAAAGATCTTG 1438

Db 182 CGGAACTCCGGCTCCCGCCTTTCGGGATCCGGCAGGCGGA-GTGGCCCCGCAAGTCACTCG 240

QY 1439 GTTTTACTACGTTATGTTTGGACGGCGATGATG 1473

Db 241 CGGTGCTACCTCGGATTTCTGGTCCGCAACCTTG 275

RESULT 6
US-08-948-564-9
Sequence 9, Application US/08948564
Parent No. 6121512
GENERAL INFORMATION:
APPLICANT: Slonimsky, Balazs
APPLICANT: Dewey, Ralph E.
APPLICANT: Cobbin, Frederick T.
TITLE OF INVENTION: No. 6121512 Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:

NAME/KEY:	CDS
LOCATION:	6..1601
US-08-948-564-9	

Query Match	5.9%	Score 93.8	DB 3	Length 1788
Best Local Similarity	48.0%	Pred. No. 4.6e-22		
Matches 303; Conservative	0	Mismatches 322	Indels 6	Gaps 1

QY	933	TTTCAGATAAATTATCCAGCCCGGACATATATCCGCGTCTTTGGAGATGATATTCAGAG	932
Db	932	TTTCGGTATATATCATATACCATCATATCAAGGCTACTTGTGATATGTGATTTTAAAGG	991
QY	993	AACAGACAGAGTTGGCGTCTTAATCGAGTGAATCCTGCTAGAGATGGTCTTATCCAGA	1052
Db	992	AAGGACACCCACCATGATTTTCACTTAATAGGTGCTATCTCTGTACTTAACATCAAT	1051
QY	1053	TATGCATTCACCGGTACAAACGACGTGATCAAGTATGTGGGAAATCAGAGCCCTAGA	1112
Db	1052	GGAATCTAAAAAAGTCCAAAGATGATGGACCTTATATTTGGGAAGGACAGAAAGTTGA	1111
QY	1113	TGAATCTGACTTGGCTTCACCTTCATATTCATACGCGCTGTGGTGAAGAAGTATTGAGCT	1172
Db	1112	AGAACTCTACATTAACAAAGTTGGTGTACCTCAGGCCATTTGTGAAGAAACAATGCGGCT	1171
QY	1173	TCATCTTCACAGCCCATCTTATATATGAGGCCCGTTTGGCCATTAACAACGATCGTTGA	1232
Db	1172	GTATTCACCAAGTCTCTTATCACCCCTTCTGTGACCATGGAACATGCACTTCTCAGG	1231
QY	1233	TGATGATTTATTCGCGACAGGAGACACAGCATGTGAACATGTGGCCGATATGCAATGA	1292
Db	1232	TGGCATACAAATCCTGTGCGGACACGTTTAAATGTGAATGCTTGAAGATCCACCGGGA	1291
QY	1293	TCACACATGTGGGTTTATCTTTTGAATTAAACCTGAGAGGTTCGTGGCAAAAGAGG	1352
Db	1292	TGATGTATTTTGAATGATCTCTCATGATTTAAGCGCTGAAGAGTTCTTGAACAGCCACA	1351
QY	1353	TGAGGTGAGTTCCTCGTTCCTTGGGTCGGAATTGAGACTTGACACTTTCGSGTGGGTGC	1412
Db	1352	AGATGTATGATGGAAGGGCTCAGAACTAATGAG-----CTCCTCCCTTTTGGTTCGGAAG	1405
QY	1413	TCGGAATTTGCCCCCGGAGAAATCTTGGTTCCTTAACACGTTATGTTTGGAGCGCGATGAT	1472
Db	1406	GAGAGCATGCCCCGTGAGCGCTCGCTGGCTCTCGATGTGGTGAACCTTGAACCAAGGGTAACT	1465
QY	1473	GTTACATGAGTTTGAATGGGAGCGTCCGATGTGAACGGCGTTGACTTATCTGAAAACT	1533
Db	1466	GTTACATCTTTCAATGTGTGCTTCTCCTTCAAACTCAAGTTGTGACATGACAGAGAGCAT	1525
QY	1533	GAGGCTTTCTTCGAGATGGCTAATCCTCTT	1563
Db	1526	TGGAATCAAAATTTAAAGCAACCCGCTT	1556

RESULT 7
 US-08-532-0658-1
 , Sequence 1, Application US/085320658
 , Patent No. 5753507
 , GENERAL INFORMATION:
 , APPLICANT: Ono, Takeshi
 , APPLICANT: Mizutani, Masaharu
 , TITLE OF INVENTION: Plant Genaciol/Nerol 10-Hydroxylase and
 , TITLE OF INVENTION: DNA Coding Therefor
 , NUMBER OF SEQUENCES: 6
 , CORRESPONDENCE ADDRESSES:
 , ADDRESSEE: No. 5753507artis Corporation
 , STREET: 59 Route 10
 , CITY: East Hanover
 , STATE: NJ
 , COUNTRY: USA
 , ZIP: 07936
 , COMPUTER READABLE FORM:
 , MEDIUM TYPE: Floppy disk
 , COMPUTER: IBM PC Compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/532,065B
FILING DATE: 22-SEP-1995
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Meigs, J. Timothy
REGISTRATION NUMBER: 38,241
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8587
TELEFAX: 919-541-8689
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 1893 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: P450-4
FEATURE:
NAME/KEY: CDS
LOCATION: 6..1490
US-08-532-065B-1

Query Match 5.9%; Score 93.6; DB 1; Length 1893;
Best Local Similarity 50.7%; Pred. No. 5,6e-22;
Matches 287; Conservative 0; Mismatches 264; Indels 15; Gaps 2;

941 AATTATCCGACCCGAGCAATATCGCGTCTTTGGAGATGATATTCAGAGAACAGACA 1000
889 AACTCTCATAGCATATGAAACACTTCTTGATATGTTTACAGACAGACAGATA 948
1001 CAGTGGCGGTCTTATTCAGAGTGGATCCTCGTAGATGATGCTTCATCCAGATATGCAT 1060
949 CAGCTCTAGTACCTCGAGTGGCCAGATGACAGAGTTCTTAAACCTTAAACAGATG 1008
1061 CAACGCTACAAACAGAGTCAATCAAGTATGCGGAAATCAAGACCTTATGATGATG 1120
1009 CAAAAGCTCAGCGCCGATCGATGATGATGATGATGATGATGATGATGATGATGATG 1068
1121 ACTTGGCTTCACTTCCATATCTTAAGCGGTGTTGAAAGATATGAGGCTTCACTCCTC 1180
1069 AATATCTCAAACTGCGTATTTTCAAGCAGTGTGTGAAGAACTTCCGTTACATACGC 1128
1181 CAGGCCACTTCTATCATGAGGCGCTTTGGCATTACAGACAGATCGTTGATGATGATG 1240
1129 CAGTTCGCTTCT---TATCCGCGAAAGCCGAATCCGATCGGAGATTTCTTGGTTTCA 1185
1241 TGTTCCTCGCAGGAGACACAGCAATGTGTAATGTGGGCGCTATTCGATTCACACG 1300
1186 TGTGTCTTAAAGTATCTCAGGTTCTAGTGAAGCTGGCCATAGACAGACCCGAGCG 1245
1301 TGTGGGTGATCTCTTGGATTAAACCTGAGGTTCTGTGCAAAAGAGGTGAGGTGG 1360
1246 TGTGGGATATCTGCTCCAGTTTGAACAGAGAGGTTTTTGGGAAAGATATGACGTGA 1305
1361 AGTTTTCGTTCTTGGGTGAGATTGAGACTTGCACCTTTCGGGTGCGGTCTGCGATT 1420
1306 GAGGTAGAGATTATGAG-----CTTACACATTCGCGCGGAGCGATGAAATTT 1353
1421 GCCCGGAGAAATCTTGGTTTACTACGTTATGTTTGAACGCGGATGATGTTACATG 1480
1354 GCCCGGAAATGCTTTGGCTATGAAAGAGTGTCTTATGCTGTGTTCTCTTCTTAT 1413
1481 AGTTTGAATGGGAGACCGTCCGATGT 1506
1414 CTTTGTACTGAAAGCTTCCGAAAGGT 1439

RESULT 8
US-09-033-055A-3
Sequence 3, Application US/09033055A
Patent No. 6069241
GENERAL INFORMATION:
APPLICANT: OHKAWA, HIDEO
APPLICANT: IMASHI, HIROMASA
TITLE OF INVENTION: NOVEL CYTOCHROME P450 GENE
NUMBER OF SEQUENCES: 17
CORRESPONDENCE ADDRESS:
ADDRESSEE: PILLSBURY, MADISON & SUTRO
STREET: 1100 NEW YORK AVENUE, N.W.
CITY: WASHINGTON
STATE: D.C.
COUNTRY: USA
ZIP: 20005
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy Disk
COMPUTER: IBM compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/033,055A
FILING DATE: MARCH 2, 1998
CLASSIFICATION: 536
ATTORNEY/AGENT INFORMATION:
NAME: PAUL E. WHITE, JR.
REGISTRATION NUMBER: 32,011
REFERENCE/DOCKET NUMBER: 9437/251563
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 861-3651
TELEFAX: (202) 822-0944
TELEX: 6714627 CUSH
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 1761
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-09-033-055A-3

Query Match 5.6%; Score 88.6; DB 3; Length 1761;
Best Local Similarity 48.8%; Pred. No. 3e-20;
Matches 300; Conservative 0; Mismatches 309; Indels 6; Gaps 2;

929 ATGATTCAGATTAATATTCGACCCGAGATTAATCGCGTCTTTGGAGATGATTTCA 988
923 ATGGAAACAAACATGATGACAGAGATTAAGCTTAATTCAGGATATGATGCTG 982
989 GAGGAAACAGACAGATTTGGGTCTTAAAGAGTGAATCCTCGTAGATGATGCTTCATC 1048
983 CAGCCAGGATACATCTGCTGACCAACGATGCGCAATGGCTTGAAGGTATTAAGCATC 1042
1049 CAGATATGATCAACGATCAAAAGAGTGAATCAAGTATGCGGAAATCAAGAGCCC 1108
1043 CAATATCTCCAGAAAGATCAAGAAATTTGATTTGTTGATGCGACCGTATGG 1102
1109 TAGATGATTTGACTTGGCTTCACTTCCATATCTTAAACGCTGTGTGAAAGATTTGA 1168
1103 TTAACGAATCCGACTTGTTCATCTCAAGTACCTTGTGTGATGATGCTGAAACATTT 1162
1169 GAGTTCATCTCCAGAGCCGCTTCTATCATGAGGCGCTTTGGCCATTAACAGACGATG 1228
1163 GAATGACCTCTGCTGCTCACTTTCT---AATTCACATGATCAATTCGAGATCA 1219
1229 TTGATGCTGCTTGTTCGCGCAGGAGCAACAGATGATGAAACATGTGGCCGTATCG 1288
1220 TCAACGCTATTAATATCCAGAAAGACAGTGTCTTCATCAACACATGATGCTTGGTA 1279
1289 ATGATCAACAGTGTGGGTGATCTTGGAGTTAAACCTGAGAG---GTTGATGCA 1345
1280 GAAACCAAGATATGGAATATATGATGATGATTTAGGCCAGAGACATTTGCTGCA 1339

QY 1346 AAGAGTGAAGTGGAGTTTCGGTCTTGGATCGGATTGAGACTTGACCTTCGGGT 1405
DB 1340 ATGAATTAAGTAGAGTTGAAATAGTCATGGGCGAGATTCAAGATTTCACATTAGTG 1399
QY 1406 CGGATCGTCGATTTGGCCCCCGGAGAGATCTTGTTTACTACCGTTATGTTTGACGG 1465
DB 1400 CTGGAAGAAAGAGTGTCTCGGTGCACCAATTAGAGTGAATTTGGTACTTATGGCTTGG 1459
QY 1466 CGATGATCTTACATGAGTTTGAATGGGACCGCTCCATGTAAACGGCTTGACTTATCG 1525
DB 1460 CTGAGATTGTTCCATCTTGTGAGTGGAGCCACCAAGATGATTAAGCGCTGAAGATATTG 1519
QY 1526 AGAAGCTGAGGCTTT 1540
DB 1520 AGACAAATTGAGCTTT 1534

RESULT 9
US-08-606-505B-63

Sequence 63, Application US/0860505B
Patent No. 6114601
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: SHIMADA, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/4
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 63 :
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
US-08-606-505B-63
Query Match 5.4%; Score 86.8; DB 3; Length 2174;
Best Local Similarity 49.7%; Pred. No. 1.5e-19;
Matches 307; Conservative 0; Mismatches 302; Indels 9; Gaps 3;

QY 956 ACATATCGCGCTTCTTGGAGATGATATTCAGAGGAACAGACAGATTGCCGCTTAA 1015
DB 972 ATATCAAGGCTCTTTATTTGAACATGTTACCGCTGTAGAGATACATCTTCAAGGCTCA 1031
QY 1016 TCGAGTGAATCTTCGTAGAGATGCTCTTCAATCCAGATATGCAATCAACGTTCAAAAG 1075
DB 1032 TAGAGTGGGCGCTGGCCGAGATGCTTAAAGATCAATCATCTTAAAGCGCCCAAG 1091
QY 1076 AGCTGATCAAGTATGCGGGAATCAAGAGCCCATAGTATGATCTTGAATCTTCACTTC 1135
DB 1092 AAATGACGCTGTATGCGCGGAGACCGGCGTTCTTGAGSAGACATATCAAGTGC 1151
QY 1136 CATATCTAAGCGCTGTGATGAAGAAAGATTAAGGCTTCAATCTTCAAGGCCCATCTTAT 1195
DB 1152 CATATCTCAAGCCATCTGCAAGAAAGATCTTCAAGAAAGCATCTTC--CAGCGCTTAA 1208
QY 1196 CATTGGCCCTTTGGCCATAACAGACAGATCGTTGATGCTGTTCCCGGAGGGA 1255
DB 1209 ATCTCCACGATCGCTCGACAGATGTAAGTAATGACACTATACCAAGGCA 1268
QY 1256 CCACAGCAATGATGACATGCGCCGATCGATGATCCACAGCTGTGGTTGATCTT 1315
DB 1269 CTAGGCTCAGGCTTAACATATGGGCTATGGAAGATCCATCTGTGGGAAATCCAA 1328
QY 1316 TGGAGTTTAACTGAGAGGTTGCTGGCAAAAGAGTGAAGTGAATTTGGCTTGG 1375
DB 1329 ATGAATTTAACCTGATGAGTATTTTGGAGAAAGAAATGCCA--AGATGATCCAGAG 1385
QY 1376 GGTGGATTTGAGACTTGACACTTTCGGTGGGCTGCGATTTGCCCGGGAAGATC 1435
DB 1386 GAATGATTTTGAAGTATGCCATTTGGAGCTGGAAGAAATTTGGCGTGAACAAGAT 1445
QY 1436 TTGGTTTACTACCGTTATGTTTGGACGGCGATGATGTTACATGATTTGAATGGGA- 1494
DB 1446 TGGGAATCTTCTAGTGAATATTTTGGAACTTTGGTGCATCTTTTGTGGAAAT 1505
QY 1495 --CGTCCGATGATGAGGCGTTGACTTATCTGAGAACTGAGGCTTCTTGCGAGATG 1552
DB 1506 TGCCATCTCTGTGATTTGAATTAATGATGATGATGCTTTTGGCTTCTGTGAGAGG 1565
QY 1553 CTAACTCTTCTGCTTA 1570
DB 1566 CAGTCCCTTGTGCTGA 1583

RESULT 10

US-09-616-990-63
Sequence 63, Application US/09616990
Patent No. 6232109
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: SHIMADA, Shigeto
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/4
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-Jul-2000

Tue Jan 20 17:43:22 2004

us-10-022-025a-2.rml

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELEPHONE NUMBER: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 63:
SEQUENCE CHARACTERISTICS:
LENGTH: 2174 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Eustoma russellianum
FEATURE:
NAME/KEY: CDS
LOCATION: 92 to 1621
IDENTIFICATION METHOD: by experiment
SEQUENCE DESCRIPTION: SEQ ID NO: 63
US-09-616-990-63

Query Match 5.4%; Score 86.8; DB 3; Length 2174;
Best Local Similarity 49.7%; Pred. No. 1.5e-19;
Matches 307; Conservative 0; Mismatches 302; Indels 9; Gaps 3;
QY 956 ACATATTCGCCGCTTTTGGAGATGATATTCAGAGAAACAGACAGCTTGGCTTTAA 1015
DB 972 ATATCAAGGCTCTTTTATGAACATGTTTACCGCTGTACGATACATCTTCAAGCTCA 1031
QY 1016 TCGAGTGAATCCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1075
DB 1032 TAGAGTGGCGCTGGCCGAGTTGCTAAGAAATCCATCATCTTAAGACAGCCCAAG 1091
QY 1076 AGCTGATCAAGTATCGGGAATCAAGAGCCCTAGTGAATGATGATGATGATGATGAT 1135
DB 1092 AATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1151
QY 1136 CATATCTCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1195
DB 1152 CATATCTCAAGCTGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1208
QY 1196 CATGGGCCGCTTTGGCCATTAACAGACAGATGATGATGATGATGATGATGATGATGAT 1255
DB 1209 ATCTCCACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1268
QY 1256 CCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1315
DB 1269 CTAAGCTCAGCTTAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1328
QY 1316 TGAAGTTAAACCTGAGAGGTTCTGAGCAAAAGAGTGAAGTGAAGTGAAGTGAAGTGA 1375
DB 1329 ATGATTTAAACCTGATAGGTTTGAACGAAAGATGCA--AGATGATGATGATGATGAT 1385
QY 1376 GGTGAGTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1435
DB 1386 GAAATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1445
QY 1436 TGGGTTTAACTACCGTTATGTTTGAACGCGATGATGATGATGATGATGATGATGATGAT 1494
DB 1446 TGGGAATACCTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1505
QY 1495 --CCGTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1552
DB 1506 TCCCATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1565
QY 1553 CTAATCCTCTCTGCTGA 1570
DB 1566 CAGTGCCTTGTGCTGA 1583

RESULT 11
US-08-606-505B-64
Sequence 64, Application US/08606505B
Patent No. 6114601
GENERAL INFORMATION:
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: KIKUCHI, Yasuhiro
APPLICANT: SHIMADA, Yukihisa
APPLICANT: OHBAYASHI, Masaya
APPLICANT: SHIMADA, Ritsuko
APPLICANT: OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSEE: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.50 inch, 720 kb storage.
COMPUTER: IBM PS/4
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/606,505B
FILING DATE: 23-MAR-1996
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992
ATTORNEY/AGENT INFORMATION:
NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Campanula medium
FEATURE:
NAME/KEY: CDS
LOCATION: 180 to 1748
IDENTIFICATION METHOD: by experiment
US-08-606-505B-64
Query Match 5.4%; Score 86.2; DB 3; Length 1927;
Best Local Similarity 49.9%; Pred. No. 2.3e-19;
Matches 273; Conservative 0; Mismatches 268; Indels 6; Gaps 2;
QY 959 TAATGCCGCTCTTGGGAGATGATATTCAGAGAAACAGACAGATGATGATGATGATGAT 1018
DB 1102 TTAAGCAGCTCTTGTGATTTATTCACGCGGCGACGATACATCATTAATGATGATG 1161
QY 1019 AGTGAATCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1078
DB 1162 AATGGGACATGAGCGAAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1221
QY 1079 TGAATCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1138
DB 1222 TGAACCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1281
QY 1139 ATCTAACGCTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1198
DB 1282 ATTTCAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1338

```

QY 1199 GGGCCCGTTGGCCATTAACAGACAGATCGTTGATGTCGTCCTTTGTCGGGAGGACCA 1258
DB 1339 TCCCAAGAAATCTTCAACAGAGCATGTGAAGTGCAGATTTCAACATACCAAAAACACTA 1398
QY 1259 CAGCAATGGTGAACATGTGGGCGGTATCGCATGATCCACACGTTGGGTTGATTCCTTTG 1318
DB 1399 GACTAATAGTGAACATGTGGGCAATGAGGAGGACCTTAAAGTGGGAAATCCATTG 1458
QY 1319 AGTTTAACTGAAGGTTGTCGCAAAAGAGTGAAGTGAAGTTTGGTTGCTTGGGT 1378
DB 1459 ATTTTACTCCGGAACGTTCTTGAAGTGAAGAAAC---ACGCAAAATTTGATCCGCGAGTA 1515
QY 1379 CGAATTTGAGACTTTCACCTTTGGGCTGGGCTGGTGGATTTCCCGGGGAAATCTTG 1438
DB 1516 ATCTTTTGAATTAATCCATTGGGCTGGAGCAAGATATGTGCAAGGAGCTTGAATGG 1575
QY 1439 GTTTTACTACCGTTATGTTTGGACGCGCATGATGTTACATGATTTGAATGGGACCGT 1498
DB 1576 GAGCGGCTCGTGGATACATATTAAGTACATTGGTGCACCTATTGATTTGAAATTC 1635
QY 1499 CCGATGG 1505
DB 1636 CTGATGG 1642

```

RESULT 12

US-09-616-990-64
Sequence 64, Application US/09616990
Patent No. 6232109

GENERAL INFORMATION:

APPLICANT: KIKUCHI, Yasuhiro
KIYOKAWA, Shigeto
SHIMADA, Yukihisa
OHBAVASHI, Masaya
SHIMADA, Ritsuko
OKINAKA, Yasushi
TITLE OF INVENTION: NOVEL PLANT GENES
NUMBER OF SEQUENCES: 67
CORRESPONDENCE ADDRESS:
ADDRESSER: FITZPATRICK, CELLA, HARPER & SCINTO
STREET: 30 Rockefeller Plaza
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10112-3801

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette - 3.50 inch, 720 Kb storage.
COMPUTER: IBM PS/IV
OPERATING SYSTEM: MS-DOS Ver3.30
SOFTWARE: PATENT AID Ver1.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/616,990
FILING DATE: 14-Jul-2000

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP44963/92
FILING DATE: 02-MAR-1992

ATTORNEY/AGENT INFORMATION:

NAME: Perry, Lawrence S.
REGISTRATION NUMBER: 31865
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-218-2100
TELEFAX: 212-218-2200

INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:
LENGTH: 1927 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA to mRNA
ORIGINAL SOURCE:
ORGANISM: Campanula medium
FEATURE:
NAME/KEY: CDS

```

; LOCATION: 180 to 1748
; IDENTIFICATION METHOD: by experiment
; SEQUENCE DESCRIPTION: SEQ ID No: 64
US-09-616-990-64

```

Query Match 5.4%; Score 86.2; DB 3; Length 1927;
Best Local Similarity 49.9%; Pred. No. 2, 3e-19;
Matches 273; Conservative 0; Mismatches 268; Indels 6; Gaps 2;

```

QY 959 TAATGCGGTCCTTTGGAGATGATTAATCAAGAGACAGACACAGTTCGGTCTTAATCG 1018
DB 1102 TTAAGGCACTCTTTTGGATTATTAATCAAGGCGGAGCAGATATCAATCAAGTGTATCG 1161
QY 1019 AGTGATCTCTGATAGATGCTCTTATCATCAGATATGCAATCAAGGTACAAAGAGAGC 1078
DB 1162 AATGGGCACTAGCGGAAATGTGAACCATGACAGATCTTAACCGGAGCCGAGAGAAA 1221
QY 1079 TGCATCAAGTATGTCGGGAAATCAAGAGCCCTAGATGATCTGACTTGGCTTCACTTCAT 1138
DB 1222 TGAACCAAGTATTTGGCAGAAACAGAAAGACTAGAAACATGACATACCAACTTGCAT 1281
QY 1139 ATCTAAGGCTGTGTGTAAGAAAGATTAATGAGCTTCACTCCAGGCGCACTTATCAT 1198
DB 1282 ATTTCCAGCCATATGCAAGAAACATTCGAAACACCTTC---CACGCCCTTAACC 1338
QY 1199 GGGCCCGTTGGCCATTAACAGACAGATCGTTGATGTCGTCCTTTGTCGGGAGGACCA 1258
DB 1339 TCCCAAGAAATCTCAACAGAGCATGTGAAGTGCAGATTTCAACATACCAAAAACACTA 1398
QY 1259 CAGCAATGGTGAACATGTGGGCGGTATCGCATGATCCACACGTTGGGTTGATTCCTTTG 1318
DB 1399 GACTAATAGTGAACATGTGGGCAATGAGGAGGACCTTAAAGTGGGAAATCCATTG 1458
QY 1319 AGTTTAACTGAAGGTTGTCGCAAAAGAGTGAAGTGAAGTTTGGTTGCTTGGGT 1378
DB 1459 ATTTTACTCCGGAACGTTCTTGAAGTGAAGAAAC---ACGCAAAATTTGATCCGCGAGTA 1515
QY 1379 CGAATTTGAGACTTTCACCTTTGGGCTGGGCTGGTGGATTTCCCGGGGAAATCTTG 1438
DB 1516 ATCTTTTGAATTAATCCATTGGGCTGGAGCAAGATATGTGCAAGGAGCTTGAATGG 1575
QY 1439 GTTTTACTACCGTTATGTTTGGACGCGCATGATGTTACATGATTTGAATGGGACCGT 1498
DB 1576 GAGCGGCTCGTGGATACATATTAAGTACATTGGTGCACCTATTGATTTGAAATTC 1635
QY 1499 CCGATGG 1505
DB 1636 CTGATGG 1642

```

RESULT 13

US-08-948-564-11
Sequence 11, Application US/08948564
Patent No. 6121512

GENERAL INFORMATION:

APPLICANT: Siminszky, Balazs
APPLICANT: Dewey, Ralph E.
TITLE OF INVENTION: No. 6121512el Cytochrome P-450 Constructs and
TITLE OF INVENTION: Methods of Producing Herbicide-Resistant Transgenic Plants
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Virginia C. Bennett
STREET: PO Box 37428

SEQUENCE CHARACTERISTICS:

LENGTH: 27627
CITY: Raleigh
STATE: No. 6121512th Carolina
COUNTRY: USA
ZIP: 27627
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/948,564
FILING DATE:
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Bennett, Virginia C.
REGISTRATION NUMBER: 37,092
REFERENCE/DOCKET NUMBER: 5051-409
TELEPHONE: 919-854-1400
TELEFAX: 919-854-1401
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1657 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1548
US-08-948-564-11

Query Match 5.0%; Score 79.6; DB 3; Length 1657;
Best Local Similarity 49.1%; Pred. No. 4.3e-17;
Matches 276; Conservative 0; Mismatches 274; Indels 12; Gaps 2;

QY 956 ACATATATGCGCGCTTTGGAGATGATATTCAGAGGAACAGACAGTGGCGTCTTA 1015
DB 914 ACATTAAGCGCGTGTCAATGAACCTTTATAGAGGAACAGACCGAGTCCGACAA 973
QY 1016 TCGAGTGAATCTCGCTAGATGTCTCTTCATCCAGATGCAATCAACGCTACAAACG 1075
DB 974 TAGTTGGGCAATGAATGACATGTTGAAGATCCCAATGATGAGCAAGGTTCAAGAG 1033
QY 1076 AGCTGGATCAATGATCGGGAATTAAGACCTTAGATCTGACTGTGCTTCACTTC 1135
DB 1034 AAGTGAAGAACTTATTCGTGTGACAAAGATTCATTAAGAAATGATGTCGAAAGCCTTC 1093
QY 1136 CATATCTAAGCGCTGTGTGAAGAAAGATTAAGGCTTCATCTCCAGGCCACTTCTAT 1195
DB 1094 CTATCTCAAGGACATGTGTGAAGAGACATTAAGTTATCCCACTTCACTACT-- 1151
QY 1136 CATGGGCCCGTGTGGCCATPAACAGACGATGCTTGAATGTGCTGCTTGTCCGACAGGA 1255
DB 1152 -TTTCCCAAGGGTAACTGAATGAAATGCAATGAAAGGTTAGCAAAATTCAGCCAAA 1210
QY 1256 CCACAGCAATGATGTAACATGTGGGCCGTATGCGATGATCCACAGTGTGGTGTATCCTT 1315
DB 1211 CTATGTGATGTATATGATGATGAGGCAATAGCAAGGACCTGAGAAATTTGGGAAGAGCTG 1270
QY 1316 TGGAGTTTAACTTGAAGGTTCGTGGCAAAAGAAAGTGAAGGTGAGTTTCGGTCTTG 1375
DB 1271 AAAAAATTTTCCCCCAAGGTTCTTGAAGTTCATGAGTTTAAAGGGAATAT-- 1326
QY 1376 GGTGGAATTTGAGACTTGACCTTTCGGGTGGGGTCTCGGATTTGGCCCGGGAAGATC 1435
DB 1327 -----GAGTTTAAGTGTATCCGTTTGGTTCGGAAGAAATGTGCTTCGGAAGCA 1381
QY 1436 TTGGTTTAACTACCGTTATGTTTGAACGGCGATGATGTACATGAGTTTGAATGGGAC 1495
DB 1382 TGGGAATTTGATGTGAGCTTCTCTGCTAATCTCATTCACACGTTTGAATGGGAAG 1441
QY 1496 CGTTCGATGTAAACGGCGTTGA 1517
DB 1442 TGGCTTAAGGCTTTCAGCAAGGA 1463

RESULT 14
US-09-564-808-5
Sequence 5, Application US/09564808
Patent No. 6501004
GENERAL INFORMATION:

APPLICANT: Selvaraj, Gopalan
APPLICANT: Nair, Ramesh B
APPLICANT: Joy IV, Richard W
APPLICANT: Keller, Wilfred A
APPLICANT: Datta, Raju S
TITLE OF INVENTION: Transgenic Reduction of Sinapline in Crucifera
FILE REFERENCE: 44618 Sequence Listing
Patent No. 6501004
CURRENT APPLICATION NUMBER: US/09/564,808
CURRENT FILING DATE: 2000-05-05
PRIOR APPLICATION NUMBER: 60/132,800
PRIOR FILING DATE: 1999-05-06
PRIOR APPLICATION NUMBER: 2,270,417
PRIOR FILING DATE: 1999-05-06
NUMBER OF SEQ ID NOS: 22
SOFTWARE: Patent In Ver. 2.1
SEQ ID NO 5
LENGTH: 1835
TYPE: DNA
ORGANISM: Brassica napus
FEATURE:
NAME/KEY: CDS
LOCATION: (2)..(1540)
US-09-564-808-5

Query Match 4.9%; Score 78.6; DB 4; Length 1835;
Best Local Similarity 48.7%; Pred. No. 1e-16;
Matches 284; Conservative 0; Mismatches 284; Indels 15; Gaps 2;

QY 925 CTCATGATGCTGATTAATATCCGACCCGACATATTCGCGTCTTTGGAGATGATA 984
DB 872 CTTCAGAAATCTATCAATTAATCCGTCATATCAATATCAAGCATATGACCTCATG 931
QY 985 TTCAGAGAAACAGACAGTTCGCGTCTTAATCGAGTGAATCTCGTAGATGTCTT 1044
DB 932 TTCAGAGAAACGAAACGTTACTTCTGTAATAGATGGGCTTAAACGAGTTATTACGG 991
QY 1045 CATCGATATGCAATCAACGTTACAAACAGCTGATCAAGTATGTGGGAAATCAAGA 1104
DB 992 AGCCCCGAGATCTTAAGAGGTTCCAAACAGAACTCGTGAAGTGTGGAATGACCGA 1051
QY 1105 GCCCTGATGAATCTGACTTGGCTTCACTTCAATCTTAACGCGTGTGTGAAGAGTA 1164
DB 1052 CGGTGAAGAAATCAACATCGAAGATGATCTTTCTGTAATGACACTGAAGAAC 1111
QY 1165 TTGAGGCTTCATCTTCAGAGCCCACTTCTATCATGAGGCCCGTTGGCCATTAACAGACG 1224
DB 1112 CTAAAGTTACACCCACGATCCCACTCT--CTTCAAGAAACGCGAGAGACACT 1165
QY 1225 ATGCTGATGTGCTGTGTTGTCGGGAGGAGCAACAGCAATGTGAACTGTGGCCGTA 1284
DB 1166 GAGATGACGTTACTTCTGTTCCCAAGAAATGCGCGTTATGATCAACCGCTTGCAT 1225
QY 1285 TCGCATGATCCACAGCTGTGGTTGATCTTTGAGTTTAACTGAGAGTTTGTGCA 1344
DB 1226 GGAAGTGAACCTTAATCTTGGCTGAGCCGAAACGTTTAAACGCTGAGGTTTAA 1285
QY 1345 AAGAGAGTGAAGTGAAGTTTTCGCTTCTGAGTGGAGTTTGAACCTTCACTTCCG 1404
DB 1286 CCGGAGATGAGG-----ATTTCAGAAAGATTAATCTTGAAGTTTATCATCTCCG 1336
QY 1405 TCGAGTGTGCGATTTTGGCCCGGAGAACTTGTGTTTACTACCGTTATGTTTGGACG 1464
DB 1337 TCGGTCGTAAATCGTGGCCGAGTATGCAACTCGGTTTATACGCGTTGAGTTAGCCGT 1396
QY 1465 GCATGATGTATCATGATTTGAATGGGAGCCGTCGATGATA 1507
DB 1397 GCTCATATTAATCATGCTTTCATGAGAAATTAATCTGATGGA 1439

RESULT 15
US-09-091-432-1
Sequence 1, Application US/09091432

/ Patent No. 5981837
/ GENERAL INFORMATION:
/ APPLICANT: Chapple, Clint
/ TITLE OF INVENTION: A Method For Regulation Of Plant Lignin Composition
/ FILE REFERENCE: 7024-325
/ CURRENT APPLICATION NUMBER: US/09/091,432
/ CURRENT FILING DATE: 1998-06-18
/ EARLIER APPLICATION NUMBER: PCT/US96/20094
/ EARLIER FILING DATE: 1996-12-19
/ EARLIER APPLICATION NUMBER: US 60/009,119
/ EARLIER FILING DATE: 1995-12-22
/ EARLIER APPLICATION NUMBER: US 60/013,388
/ EARLIER FILING DATE: 1996-03-14
/ NUMBER OF SEQ ID NOS: 3
/ SOFTWARE: Microsoft Word 2.0C
/ SEQ ID NO 1
/ LENGTH: 1838
/ TYPE: DNA
/ ORGANISM: Arabidopsis thaliana
US-09-091432-1

Query Match 4.9%; Score 78.6; DB 2; Length 1838;
Best Local Similarity 48.7%; Pred. No. 1e-16;
Matches 284; Conservative 0; Mismatches 284; Indels 15; Gaps 2;
QY 925 CTCGATGCTTCAGATTAATTCGACCCGGACATATCGCCGTTCTTTGGAGATGATA 984
DB 908 CTTCATAATTCATCAACTTACCCGTGACAAATCAAGCAATCATGAGCGTTATG 967
QY 985 TTCAGAGAACAGACACAGTTGGGCTTATCGATGGATCCTGCTAGATGCTCTT 1044
DB 968 TTTCGAGAACGGAACGGAGTGGGCGATAGAGTGGCTTAAAGGATTATTACGG 1027
QY 1045 CATCCAGATATGCAATCAACGATCAAAAGAGCTGATCAAGTATGCGGAATCAAGA 1104
DB 1028 AGCCCGAGATTAATAACGGGTCACAAAGAACTGCCGAAGTCGTTGACTTGACAGA 1087
QY 1105 GCCCTAGATGATCTGACTTGGCTCACTTCATATCTAAGGCTGTGGTGAAGAAGTA 1164
DB 1088 CGAGTTGAAGAAATCCGACATCGAAGATGACTTATCTCAATGCACTCAAGAAACC 1147
QY 1165 TTGAGCTTCATCTCCAGGCCCATCTTATCATGAGGCCGTTGGCCATTAACAGACAG 1224
DB 1148 CTAAAGATGACCCACCGATCCCTCTCT-----CCTCCAGAAACCGGAGGACACT 1201
QY 1225 ATGCTGATGATGCTGCTGTTGCGGAGGAGCAACAGCAATGATGTAACATGTGGCCGTA 1284
DB 1202 AGTATGACGGTTCTTCATTTCCCAAGAAATCTGCTGATGATCAACGGCTTTGCCATA 1261
QY 1285 TCGCATGATCCACAGTGTGGTGTGATCTTTGAGTTTAAACCTGAGAGTTTGTGCA 1344
DB 1262 GGAACGCACTTGTGACTGACCCGAGACAGCTTAAACCATCGAGGTTTGGAA 1321
QY 1345 AAGAGGTGAGTGTGAGTTTGGCTTCTTGGGTGAGTTTGAACCTGACCTTTGGG 1404
DB 1322 CCGGGCGTACCG-----ATTCAAGGAGCAATTCAGTTTATACCGTTGGG 1372
QY 1405 TCGGATGCTGGAATTTGCCCCGGAAGAAATCTTGTTTAACTAACCGTTATGTTTGGACG 1464
DB 1373 TCGGATGCTGATGCTGCCCCGGGATGCACTAGGTTTATACGCGCTTACCTTACGCGTG 1432
QY 1465 GCCATGATGTTACATGAGTTTGAATGGGACCGTCCGATGTA 1507
DB 1433 GCTCATATATTAATGCTTCACGTGGAATTAACCTGATGGGA 1475

score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

SUMMARIES

11088.430 Million cell updates/sec

Sequence: 1 ataataataaaaatgta.....atatggaattagagatttc 2009

Searched: 2888711 beqs, 20454813386 residues

Total number of hits satisfying chosen parameters: 4

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing: Listing first 45 summaries

Database : GenEmbl:★

1:	gb	ba:	*
2:	gb	hig:	*
3:	gb	in:	*
4:	gb	om:	*
5:	gb	ov:	*
6:	gb	pac:	*
7:	gb	ph:	*
8:	gb	pl:	*
9:	gb	ro:	*
10:	gb	ro:	*
11:	gb	sts:	*
12:	gb	sy:	*
13:	gb	va:	*
14:	gb	va:	*
15:	em	ba:	*
16:	em	fun:	*
17:	em	hum:	*
18:	em	in:	*
19:	em	mu:	*
20:	em	om:	*
21:	em	ov:	*
22:	em	ov:	*
23:	em	pac:	*
24:	em	ph:	*
25:	em	pl:	*
26:	em	ro:	*
27:	em	sts:	*
28:	em	un:	*
29:	em	va:	*
30:	em	hig	hum:
31:	em	hig	inv:
32:	em	hig	other:
33:	em	hig	mus:
34:	em	hig	pin:
35:	em	hig	rod:
36:	em	hig	mam:
37:	em	hig	vrt:
38:	em	sv:	*
39:	em	hgo	hum:
40:	em	hgo	mus:
41:	em	hgo	other:

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	2009	100.0	55870	8	AC005819	Arabidops
c 2	2009	100.0	84835	8	AC006418	Arabidops
c 3	232	11.5	740	8	ATH552847	Arabidops
c 4	141	7.0	504	8	ATH553770	Arabidops

ALIGNMENTS

LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM
RESULT 1 AC005819/c	AC005819 Arabidopsis thaliana chromosome 2 clone T34 map C106C03, complete sequence.	AC005819	AC005819	AC005819.3	GI:20197448	HTG. Arabidopsis thaliana (thale cress) Arabidopsis thaliana

REFERENCE
AUTHORS
JOURNAL
REFERENCE
AUTHORS
JOURNAL
TITLE
JOURNAL

1 (bases 1 to 55870) Kall, S., Shea, T.P., Fujii, C.Y., Mason, T.M.,
Rounsley, S.D., Lin, X.,
Shen, M., Romling, C.M.,
Fraser, C.M., Somerville, C.R. and Venter, J.C.
Unpublished
2 (bases 1 to 55870)
Lin, X.
Direct Submission
Submitted (09-MAR-2000). The Institute for Genomic Research, 9712

REFERENCE 3 (bases 1 to 55870)
AUTHORS Town, C.D. and Kaul, S.
TITLE Direct Submission
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr. Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT On Apr 18, 2002 this sequence version replaced gi:559489.

FEATURES	Location/Qualifiers
source	1. .55870

misc_feature

gene

MURINA

CDS

Pred. No. is the number of results predicted by chance to have a

EILRSPOGKNTGGSGTHSSRVSHGTGKCRRLQSSSSSSDGEGRKTTKNVKNKN
 SNLPMKAFMNSNVQMINNSIVNSTASHHDPGVHLKISRKPSGSDNGFHVSDYGNNGY
 TN"
 repeat_region 723..810
 /rpt_family="CACCAT)n"
 repeat_region 1345..1382
 /rpt_family="AT_rich"
 repeat_region complement(1904..1925)
 /rpt_family="AT_rich"
 gene 3740..5558
 /gene="At2g46640"
 /note="synonym: T3A4.2; predicted by genSCAN and
 genefinder"
 mRNA join(43740..4072,4525..4839,5083..5215,5407..5558)
 /gene="At2g46640"
 join(3740..4072,4525..4839,5083..5215,5407..5558)
 /gene="At2g46640"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAC69921.1"
 /db_xref="GI:3831438"
 /translation="MWKELIDVGRNKKKKRNEGISIEKTKAILDQGLVDALDWF
 DGVLTITIGFSDTLNFKEDMDDECGSDLDLYVIDGIIINVQESDPLISVKNK
 FYDHEVDVGAIDTHFGSIKTAETPAVTAALAEI EPKRTTLAELFMDHDGDTYD
 HCKKKNPNLDGEVKKYKONGSKLSRKFSEVKKLVMSKKEENDLPIKKRQMI
 KRMKKKIHDPDVKALKKDVPRKPRNCEALSHLKIQCVGIVMTVEMYCYLL
 TNITKLVFVPMHIVFYVVICLTITYNASF"
 gene 6216..7156
 /gene="At2g46650"
 /note="synonym: T3A4.3; supported by full length cDNA;
 Ceres:3743"
 join(6216..6447,6552..6618,6704..7156)
 /gene="At2g46650"
 join(6370..6447,6552..6618,6704..6957)
 /gene="At2g46650"
 /codon_start=1
 /product="putative cytochrome b5"
 /protein_id="AAC69922.1"
 /db_xref="GI:3831439"
 /translation="MANLISFHDVAKKKNKNDIILHGKYVDISTFMDHFGDQVNL
 LAVGKQASIDFEDVNSKQAKELMKKCYCTGDVQSTVPTQQYIPWMEKESVAAET
 KEBSGKLLITLPLILIGVAFALRFNNK"
 repeat_region complement(7227..7247)
 /rpt_family="AT_rich"
 repeat_region 7693..7715
 /rpt_family="AT_rich"
 repeat_region complement(7799..7832)
 /rpt_family="AT_rich"
 gene complement(8165..9980)
 /gene="At2g46660"
 /note="synonym: T3A4.4"
 complement(join(8165..8782,9006..9980))
 CDS complement(join(8165..8782,9006..9980))
 /gene="At2g46660"
 /codon_start=1
 /product="putative cytochrome P450"
 /protein_id="AAC69923.1"
 /db_xref="GI:3831440"
 /translation="MATKLBSLIPALSKCSVLSQTNLAFSLAVTIIWLALSFLM
 TYPGAPAMKYLFGRLISGYKGVNIPGKGFPLVSGMSLMSSTLARRADAEK
 GAKRLAPSLIGETRIYVTCNPVAKELINSPEVADRPKESAGLMEFRAIGFAPHY
 VYRTLRRTASNLHSTKQIRRAETORRIVISOMVEFLKOSNBPFCYBELTKASIN
 NMCSVPGQEVYIEKQHYELPMDEGVILGTNMTDHLPLSEPPORLRSRCTL
 VKVNRVSRITISEHRNOTGDPDPFVDVLSSHSDELSPDILATIMEMIFRGCT
 VAVILWILARVHLPMOSTVQNSLSDVSGSLSDSDSLASPLIFLAVKELRLH
 PEGPLISVARLAIPTDITVDGLVPAQTAMNMAVSHDPHWDPLEFKERRVAKE
 GEVSFVLSGDLRLAPFGSGRI CPKXNIGFTTVVFTWAMHLFEWQPSDNGVDLS
 EKLRLSCMANLPAKLRRRS"
 repeat_region 10124..10180
 /rpt_family="AT_rich"
 repeat_region complement(14979..15002)
 /rpt_family="(A)n"

repeat_region 15546..15568
 /rpt_family="AT_rich"
 repeat_region 17214..17366
 /rpt_family="AT_rich"
 repeat_region 17478..17503
 /rpt_family="(A)n"
 repeat_region complement(17523..17543)
 /rpt_family="AT_rich"
 repeat_region complement(17647..17688)
 /rpt_family="AT_rich"
 repeat_region complement(17874..17894)
 /rpt_family="AT_rich"
 repeat_region complement(18001..18060)
 /rpt_family="AT_rich"
 gene complement(19152..19796)
 /gene="At2g46670"
 /note="synonym: T3A4.5; predicted by genSCAN and
 genefinder"
 mRNA complement(join(19152..19310,19344..19796))
 /gene="At2g46670"
 complement(join(19152..19310,19344..19796))
 /gene="At2g46670"
 /codon_start=1
 /product="hypothetical protein"
 /protein_id="AAC69924.1"
 /db_xref="GI:3831441"
 /translation="MYLTCRFEESKSAKAVVALESTGSEPKTPTESHEKLRVSD
 QSSSTTSNQNIGSSSVSFNVOVLQSTVYQKQSDSPIPVESNREKASKEVSGS
 TNEGIAGQSGSTKREKESAKORRSORALPMRLKRDRCFQKLIKQIQDOT
 DSMWYKGLQVRYSQRKLAQRPRVKGQVFRVNSDASTKS"
 gene 19989..21343
 /note="molecular marker Atb7"
 complement(20037..21512)
 /gene="At2g46680"
 /note="synonym: T3A4.6; identical to SP:146897; supported
 by cDNA: gi15027938.gb AY045826.1"
 complement(join(20037..20039,20111..20756,20977..21512))
 /gene="At2g46680"
 complement(join(20340..20756,20977..21336))
 /gene="At2g46680"
 /codon_start=1
 /product="homeodomain transcription factor (ATHB-7)"
 /protein_id="AAC69925.1"
 /db_xref="GI:3831442"
 /translation="WTGGGYSPAMSAEPLITMKMKKSNHNKNORFDEQIKSL
 EMVFSETRLEPRKRYQIAREIGLOPROVAILMFOKRAWRKSKOLETYNLIROYND
 IASQFSLKKEQALVSELQRIKKAQKTOBEERQCGDDAVVALSSTHSEERN
 RRRKPEVAPMEMKDDKGGHGVCDHHDVDDDDGYSNNIRRETFGGFEEBPHLM
 IVEPADSCITSSDDWRGFSKDTTLTLDSSNNYPWRDWS"
 repeat_region complement(22636..22664)
 /rpt_family="AT_rich"
 repeat_region 23532..23571
 /rpt_family="AT_rich"
 repeat_region 23760..23783
 /rpt_family="AT_rich"
 repeat_region 24595..24656
 /rpt_family="(CAAAA)n"
 repeat_region complement(25135..25155)
 /rpt_family="AT_rich"
 repeat_region 25903..25923
 /rpt_family="AT_rich"
 repeat_region complement(26667..26771)
 /rpt_family="AT_rich"
 repeat_region complement(26830..26901)
 /rpt_family="Rf:ATREP1|ATREP1 ATREP1 dispersed repeat from
 GRI166608 - a consensus."
 repeat_region 30015..30038
 /rpt_family="(TAAA)n"
 repeat_region complement(31911..31933)
 /rpt_family="AT_rich"
 repeat_region complement(32444..32466)
 /rpt_family="AT_rich"
 gene 35276..35832

complete sequence.

ACCESSION AC006418
VERSION AC006418.4 GI:20197771

KEYWORDS HTG.
SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE 1 (bases 1 to 84825)
Lin, X., Kaul, S., Shee, T.P., Fujii, C.Y., Shen, M., VanAken, S.E.,
Barnstead, M.E., Mason, T.W., Bowman, C.L., Rong, C.M.,
Benito, M.-I., Carrera, A.V., Creasy, T.H., Buell, C.R., Town, C.D.,
Nierman, W.C., Fraser, C.M. and Venter, J.C.

REFERENCE 2 (bases 1 to 84825)
Unpublished

JOURNAL
AUTHORS Lin, X.

REFERENCE
JOURNAL Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA

REFERENCE
JOURNAL 3 (bases 1 to 84825)
Town, C.D. and Kaul, S.

REFERENCE
JOURNAL Submitted (27-FEB-2002) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA, cdrom@tigr.org
On Apr 18, 2002 this sequence version replaced gi:6598561.

COMMENT
FEATURES
source
1..84825
Location/Qualifiers

/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/culivar="Columbia"

/db_xref="taxon:3702"
/chromosome="2"
/map="C106C03"

/clone="F13A10"

1..36153
/note="overlap with BAC clone F11C10
(AC006526:59064..95214)."

1..36153
/note="overlap with BAC clone F11C10
(AC006526:59064..95416)."

complement(1047..3029)
/gene="At2g46480"

/note="synonym: F13A10.1; predicted by genSCAN"
complement(join(1047..1332,1409..1518,1595..1997,
2074..2431,2503..2828,2916..>3029))

/gene="At2g46480"
complement(join(1047..1332,1409..1518,1595..1997,
2074..2431,2503..2828,2916..3029))

2074..2431,2503..2828,2916..3029))

/gene="At2g46480"
/codon_start=1

/product="hypothetical protein"

/protein_id="AAD20159.1"

/db_xref="GI:4415929"

/translation="MTDACCCKGNDKRVPRFGWTIGKAFNDTPMTHSRISROK
RLERANLMDNDLQKLETAAMARSYSVSAFLNNTYTWKREYRGKSFEDMLRMD
QIMARVYSGLAKPTNNIALHOEIEIQMLKAMEBESDIDQEVLDSDRWQIIA
RAHEQLVECKLVNTKALMLQTEDELENEQYITFLTQASKALPDALHCLNRLNL
EYHLLEPMNPRENLENPELXAYALFSDNVAASVYVMAQDSRHFVHLV
TDKLNFGAMSWFELNPGELATIHVQRPEDFTNLASYSVPLSQLESAMKKFYKTA
ASBVSSESENTKYRYPKMGMTNHLRYPRIPIPKLEKLFDDDVVVOGDLTP
IDLKRVENEDPKFCGMYGKMTFDLKEKKNNTETTYHFWQMLNENRTIMKGLP
PELLTFYNDLPQKRWHLGIGDKGIDVKKIESRAVIHNGMKRTMGISKYP
WTKYTNDFHPIFCRLFE"

4060..4696
/note="At2g46490"

/note="synonym: F13A10.2"

4060..4696
/gene="At2g46490"

4100..4504
/gene="At2g46490"

/codon_start=1
/product="unknown protein"

repeat_region

repeat_region

gene

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

repeat_region

/protein_id="AAD20160.1"

/db_xref="GI:4415930"
/translation="MADGIFRSIFPGICISGLDSALERRPYRHNCCALHDKSGAK
NONORRPPSCRRHGSSEISFPIRSMSEBNTIAMNLFSSSSSSNLDQSLSSSSLSN
LASDLPVDNATEPSSRNKQLMTIDEGDD"

complement(4483..4517)
/rpt_family="CAT"n

complement(6756..6792)
/rpt_family="AT-rich"n

complement(11207..13602)
/gene="At2g46500"

/note="synonym: F13A10.3; similarity to ubiquitin family
of proteins; supported by cDNA: gi_16930424_gb_Af419566
1_Af419566"

complement(join(11207..11764,11858..>13602))
/gene="At2g46500"

complement(join(11384..11764,11858..13177))
/gene="At2g46500"

/codon_start=1
/product="expressed protein"

/protein_id="AAD20161.1"

/db_xref="GI:4415931"

/translation="MSAGVALSPVSEPLIMPLVANSGLDVPDPTIMVLTLPDS
VIPRVEDSDIESVKTIOSTRTFRANQKLVFGSELRNSNMDVAVSEGNILH
LVKLSDQVLDVTTGCKHCFRFEVERGNNGIVKQISKRGDPFUPDDQELIYEGE
KLEDOSLINDICRNDSDVHLVRSARVAVKVPKAFELSTVAPQKDKGKRAKSI
VPEPKLSLEPVVNSKAKVPLVVDMTQASDGLKSGNSVRSSEGGVAVFMDGFSN
NKPVGFPEPDEEMAEENPQGLPLSNGESLKGKGTGSGALEVAVYILDMPSN
KMGEEIGFAGVGPPTAMIEGLHFGNPKIKKISGLQGTENDSCEDMPLSP
VBRKIVSLDIRANADRRGNTNIDKSGKLVYPIHIDGTFSECECTEMY
WQAKRPSYSAETQSTIRSLDEEDIDILKFGHGMCPMPATQTLRISMLIKVVERGL
TAFETIGTCRETSLKSLVEEVEAEQEAVALPQTSALFALSLDVMIDHDSQEH"

complement(14394..15192)
/gene="At2g46505"

/note="synonym: F13A10.20; supported by full length cDNA:
Ceres: 11006"

complement(<14394..>15192)
/gene="At2g46505"

complement(14656..15111)
/gene="At2g46505"

/codon_start=1
/product="expressed protein"

/protein_id="AAM15243.1"

/db_xref="GI:20197780"

/translation="MSLRRTILDLRQOTATLSKSLPFSMSHISASATAYANPLG
RDLSSIPPAQAKLKPDSTNLVTDKSISSSISQSLNPAKFSQSSRGYTNQFLR
KIPVFIHIEGMEELIADYVQEMTRNLIWNSLGFQIIVAKDIIIFLL"

complement(15830..17530)
/gene="At2g46510"

/note="synonym: F13A10.4"

complement(<15830..>17530)
/gene="At2g46510"

complement(15830..17530)
/gene="At2g46510"

/codon_start=1
/product="putative bHLH transcription factor"

/protein_id="AAD20162.2"

/db_xref="GI:20197775"

/translation="NNMSDLGNDDEKSVAVAGHLASDFIRANSNONLFIWGT
DITLNKSLSDVDPNSENFSNNVAFVQQTSSRQQLVGLGDCCEPRNDEESKY
VRSYNNNGAEEETWODKREKVLQKHLFGSGDEVDVLSLEKVTATEIFFLASNY
FEPNHEGGPGRGYSQSKVMDVANSSEDFCSFVAKSGITIVWPTDAGVLE
LSGWSLEPNIGIVSVQVLPFRVYQPPVYVNNMGIGIKLGGDLSGAAVPEK
LEVRNLDREPTPOSWEYNNKGPFGTGPORDVYKLVENVMVDNNNYTQIEFA
GSSVANSNPSTNTQOKESGCTEKRPKSLAGAGIVVDEKRPKRGKANGSE
PLNHVAERQRRKLNQRTFALRSVYVNSKDKXSLGDATSYIKELQEKYKME
RVGTDKSLSESNITVESEVVDIAMNEBVAVRVTSPLDSPARSIIDAMNSVDE
MEKSLIADETWTFHFIKSNNGSDPLTEKLIAPVPTSTQPLPSSSQVSGDI"

17920..17946
/rpt_family="(GA)n"

complement(18392..18454)

[illegible]

JOURNAL	Submitted (21-NOV-2002) Balzergue S., IMAGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE					
COMMENT	PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbgap.verisalles.inra.fr/publiclines/ . This sequence has been generated in the framework of the French plant genomes program 'genoplante' (http://www.genoplante.com and http://genoplante-info.infobios.fr).					
FEATURES						
source	Location/Qualifiers 1..740 /organism="Arabidopsis thaliana" /mol_type="Genomic DNA" /cultivar="MaselliLewskiJa" /db_xref="taxon:3702" /clone="J345B11" /clone_id="Arabidopsis thaliana T-DNA insertion lines" misc_feature 1..740 /note="T-DNA flanking sequence left border"					
BASE COUNT	192 a	172 c	197 g	179 t		
ORIGIN						
Query Match	11.5% Score 232, DB 8; Length 740;					
Best Local Similarity	99.4%; Pred. No. 2..2e-117;					
Matches 332; Conservative	0;	Mismatched	2;	Indels	0;	Gaps 0;
Oy	360	CCGGAACGGTATTCGCCGCCTCAAAAGCGCTCCTTGTTGGTAGCATGTCACTCATGT	419			
Dd	423	CCGGAAACGTTATTCGCCGCCTCAAAAGCGCTCCTTGTTGGTAGCATGTCACTCATGT	370			
Oy	420	CAGAAGCTCTAAGCTCACCGAAGATGCGTGAATCGAGCAAGAGGC	479			
Dd	369	CAGAAGCTCTAAGCTCACCGAAGATGCGTGAATCGAGCAAGAGGC	310			
Oy	480	TGATGCGTTTACGCTTGAAGAGACTGCGTGAATCGTCAAGCGATCCGACGTAAGGA	539			
Dd	303	TGATGCGTTTACGCTTGAAGAGACTGCGTGAATCGTCAAGCGATCCGACGTAAGGA	250			
Oy	540	AAGAGATTGTGAATAGCCCGGTTTTGCTGATGACCGGTTAAAGATCGGTTACTAC	599			
Dd	249	AAGAGATTGTGAATAGCCCGGTTTTGCTGATGACCGGTTAAAGATCGGTTACTAC	190			
Oy	600	TGATGTTTAAAGAGCAATTTGTTTSCACAACAAGGTTTACTGCGAAGCGTTTCCG	659			
Dd	189	TGATGTTTAAAGAGCAATTTGTTTSCACAACAAGGTTTACTGCGAAGCGTTTCCG	130			
Oy	660	GTTTCGCTTCGAACCATCTCTTTTGTCATAAACA	693			
Dd	129	GTTTCGCTTCGAACCATCTCTTTTGTCATAAACA	96			
RESULT 4	ATH553770.o					
LOCUS	ATH553770 504 bp DNA linear PLN 29-MAR-2003					
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, left border, clone 371B10.					
ACCESSION	AJ553770					
VERSION	AJ553770.1 GI:29370237					
KEYWORDS	left border; T-DNA flanking sequence.					
SOURCE	Arabidopsis thaliana (chate crees)					
ORGANISM	Arabidopsis thaliana					
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; Fossils; euroside II; Brassicales; Brassicaceae; Arabidopsids.					
REFERENCE	Brunaud,V., Balzergue,S., Dubreucq,B., Aubourg,S., Samson,F., Chauvin,S., Bechtold,N., Craud,C., DeRose,R., Pelletier,G., Lepointec,L., Caboche,M. and Leclercq,A.					

TITLE T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
 JOURNAL EMBO Rep. 3 (12), 1152-1157 (2002)
 MEDLINE 22363535
 PUBMED 12446565
 REFERENCE 2 (bases 1 to 504)
 AUTHORS Balzerque, S.
 TITLE Direct Submission
 JOURNAL Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue Gaston Cremieux, 91057 Evry cedex, FRANCE

COMMENT PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.infobiogen.fr>).

FEATURES
 source 1..504
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /cultivar="Massiliaeskija"
 /db_xref="taxon:3702"
 /clone_id="371B10"
 /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
 Location/Qualifiers

misc_feature 1..504
 /note="T-DNA flanking sequence
 left border"

BASE COUNT 180 a 128 c 71 g 125 t
 ORIGIN

Query Match 7.0%; Score 141; DB 8; Length 504;
 Best Local Similarity 99.5%; Pred. No. 1.7e-66;
 Matches 191; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1714 AGAGTTCTCGTGGCAAAAGAGTGAAGTTTCGGTCTTGAGTTCGATTGAGAC 1773
DB 497 AGAGTTCTCGTGGCAAAAGAGTGAAGTTTCGGTCTTGAGTTCGATTGAGAC 438
QY 1774 TTGCACCTTCGGGTCGGGTCGTCGATTTGCCCGGAGAATCTTGTTTACTACCG 1833
DB 437 TTGCACCTTCGGGTCGGGTCGTCGATTTGCCCGGAGAATCTTGTTTACTACCG 378
QY 1834 TTATGTTTGGACGGCGATGATGTACATGAGTTTGAATGGGACCGTCCGATGTTACG 1893
DB 377 TTATGTTTGGACGGCGATGATGTACATGAGTTTGAATGGGACCGTCCGATGTTACG 318
QY 1894 GCGTTGACTTAT 1905
DB 317 GCGTTGACTTAT 306

```

Search completed: January 16, 2004, 18:05:05
 Job time : 7412 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2004, 09:30:00 ; Search time 4825 Seconds
(without alignments)

10119.729 Million cell updates/sec

Title: US-10-022-025A-1
Perfect score: 2009
Sequence: 1 aataataataataatgta.....atcgaagaattagatttc 2009

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues
Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: em_escba:*
2: em_escbun:*
3: em_escin:*
4: em_escmu:*
5: em_escov:*
6: em_escpl:*
7: em_escro:*
8: em_escr:*
9: gb_esc1:*
10: gb_esc2:*
11: gb_esc3:*
12: gb_esc4:*
13: gb_esc5:*
14: gb_esc6:*
15: em_escf:*
16: em_escf:*
17: em_escf:*
18: em_escf:*
19: em_escf:*
20: em_escf:*
21: em_escf:*
22: em_escf:*
23: em_escf:*
24: em_escf:*
25: em_escf:*
26: em_escf:*
27: em_escf:*
28: gb_esc1:*
29: gb_esc2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	DB ID	Description
1	566.4	28.2	29	BZ428523 BONHW74TF
2	550.6	27.4	28	BH492184 BOHFJ050TR
3	502.6	25.0	29	BZ428529 BONHW74TR
4	479	23.8	28	BH451907 BOGSO44TR

Result No.	Score	Query Match Length	DB ID	Description
5	472	23.5	28	BH431683 BOHF746TR
6	458.8	22.8	28	BH510149 BOHQ049TR
7	421.6	21.0	28	BH527239 BOHB061TR
8	419.6	20.9	29	BZ466758 BONND11TR
9	396.6	19.7	29	BZ443616 BONFE28TR
10	382.6	19.0	28	BH451899 BOGSO44TF
11	337.6	16.8	28	BH777803 Cde14a11.
12	330.8	16.5	28	BH334160 BOHC239TF
13	311.2	15.5	28	CB344213 CA48EN000
14	304	15.1	28	CB344213 CA48EN000
15	289.2	14.4	28	BH581933 BOHRM09TR
16	288.6	14.1	28	BH581933 BOHRM09TR
17	283.4	14.1	28	BH581933 BOHRM09TR
18	279.8	13.5	28	BH581933 BOHRM09TR
19	271.2	13.5	28	BH581933 BOHRM09TR
20	254.6	12.7	28	BH581933 BOHRM09TR
21	252	12.5	28	BH581933 BOHRM09TR
22	251.8	12.5	28	BH581933 BOHRM09TR
23	238.8	11.9	28	BH581933 BOHRM09TR
24	238.4	11.9	28	BH581933 BOHRM09TR
25	236.6	11.8	28	BH581933 BOHRM09TR
26	233.6	11.6	28	BH581933 BOHRM09TR
27	233.2	11.6	28	BH581933 BOHRM09TR
28	231.2	11.5	28	BH581933 BOHRM09TR
29	224.4	11.2	28	BH581933 BOHRM09TR
30	220.4	11.0	28	BH581933 BOHRM09TR
31	217.6	10.8	28	BH581933 BOHRM09TR
32	213.2	10.6	28	BH581933 BOHRM09TR
33	212.6	10.6	28	BH581933 BOHRM09TR
34	212	10.6	28	BH581933 BOHRM09TR
35	210.6	10.5	28	BH581933 BOHRM09TR
36	208.6	10.4	28	BH581933 BOHRM09TR
37	207.4	10.3	28	BH581933 BOHRM09TR
38	204.6	10.2	28	BH581933 BOHRM09TR
39	201.6	10.0	28	BH581933 BOHRM09TR
40	199.6	9.9	28	BH581933 BOHRM09TR
41	198.8	9.9	28	BH581933 BOHRM09TR
42	196.8	9.8	28	BH581933 BOHRM09TR
43	193.2	9.6	28	BH581933 BOHRM09TR
44	189.6	9.4	28	BH581933 BOHRM09TR
45	189	9.4	28	BH581933 BOHRM09TR

ALIGNMENTS

RESULT 1
LOCUS BZ428523
DEFINITION BONHW74TF BO.1.6.2 KB, lot Brassica oleracea genomic clone BONHW74.
GENOMIC SURVEY SEQUENCE.
ACCESSION BZ428523
VERSION BZ428523.1 GI:26671493
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Brassicaceae
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids / euroside II; Brassicales; Brassicaceae; Brassica.
REFERENCE 1 (bases 1 to 712)
TOWN, C.D., Van Aken, S., Uteback, T., Koo, H., and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BONHW74TR
CONTACT: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: 1P
Class: sheared ends.

FEATURES

Source

Location/Qualifiers

1..712
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BONHW74"
/clone_lib="BO_1.6.2 KB tot"
/note="Vector: pHOSt; Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into pHOSt using BstXI linkers"

BASE COUNT

168 a 179 c 177 g 188 t

ORIGIN

28.2%; Score 566.4; DB 29; Length 712;

Query Match Best Local Similarity 87.2%; Pred. No. 1.2e-102;
Matches 621; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 411 CACTCATGTCAAGACCTCTGACCTGACCGAATCGCTGATGCACTGAGAAATTGGAG 470
Db 1 CTCCTCATGTCCAGACCTCTGACCTCTCCGCGAATCGGACGACTCTGCGAGAGATTGGAG 60
QY 471 CCAAGAGAGGCTTTCAGCTTTCAGTACGAGAGACTGCGGATCGTACGTCGAAATCCCG 530
Db 61 CCAAGAGGCTTTCAGCTTTCAGTACGAGAGACTGCGGATCGTACGTCGAAATCCCG 120
QY 531 ACGTAGCGAAGAGATTCTGATAGCCGGTTTGTGTGATCGACCGGTTAAAGATCGG 590
Db 121 ACGTTGCTAAGAGATTCTGATAGCCGGTTTGTGTGATCGACCGGTTAAAGATCGG 180
QY 591 CTTACTCATGTATTTTAAAGACGAAATGTTTGTGACACACAGGTTTAACTGGCGAA 650
Db 181 CTTACTCATGTATTTTAAAGACGAAATGTTTGTGACACACAGGTTTAACTGGCGAA 240
QY 651 CGCTTCGCGGATGCTTCGACCACTCTTTAGTACAAACAAATCAGAAAGACCGAGA 710
Db 241 CTCGCGGCGATGCTTCGACCACTCTTTAGTACAAACAAATCAGAAAGACCGAGA 300
QY 711 CGCAACGAGAGTATCTCAAGCCGAGATGTTGATTTCTTGAAAAAGATGATAG 770
Db 301 CACAGAGGCGAGATGCTTCGACCACTCTTTAGTACAAACAAATCAGAAAGACCGAGA 360
QY 771 AACCTGTTTGTGCTGATGCTTAAACGCGCTGCTTACAAACATGATGCTCTG 830
Db 361 GCGTCTGTTTGTGCTGATGCTTAAACGCGCTGCTTACAAACATGATGCTCTG 420
QY 831 TATTGGAACAGATGATGAGCTTGAACAAAACATGTTGAGTTACGTGAATGTCGAG 890
Db 421 TTTTGGGACAGAGATGATGAGCTTGAACAAAACATGTTGAGTTACGTGAATGTCGAG 480
QY 891 AAGGTTAGATTGCTCGGAAAGCTGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 950
Db 481 AAGGTTAGATTGCTCGGAAAGCTGGAATGGAATGGAATGGAATGGAATGGAATGGAAT 540
QY 951 TTGATCTCAAGAGACTCGGCTGATGATGTTCCACTGTCGACAAAGTTAAACCGGTTG 1010
Db 541 TTGATCTCAAGAGACTCGGCTGATGATGTTCCACTGTCGACAAAGTTAAACCGGTTG 600
QY 1011 TATCCCGGATTATATCCGAGACCGTAAATCAACCGGATGATTGCTGATGATTGCTG 1070
Db 601 TATCCCGGATTATATCCGAGACCGTAAATCAACCGGATGATTGCTGATGATTGCTG 660
QY 1071 ACGTTTGTCTCCCTCCATGTTTCAAGTAAATTAATCCGACCGGACATTAAT 1122
Db 661 ACGTTTGTCTCCCTCCATGTTTCAAGTAAATTAATCCGACCGGACATTAAT 712

RESULT 2

BH492184 784 bp DNA linear GSS 13-DEC-2001
LOCUS BH492184
DEFINITION BOHFJ05TR BOHF Brassica oleracea genomic clone BOHFJ05, genomic
survey sequence.

ACCESSION BH492184
VERSION BH492184.1 GI:17700288

KEYWORDS

SOURCE

ORGANISM

GSS.
Brassica oleracea
Brassica oleracea
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
/ eurosids II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 784)
Town, C.D., Van Aken, S., Utecherback, T., Koo, H. and Fraser, C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished
Other GSSs: BOHFJ05TR
Contact: Chris Town

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

TIGR

9712

Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Classes: sheared ends.

FEATURES

source

1..784
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T01000DH3"
/db_xref="taxon:3712"
/clone="BOHFJ05"
/clone_lib="BOHF"
/note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
genomic DNA inserted into pHOSt using BstXI linkers"

BASE COUNT

204 a 189 c 211 g 180 t

ORIGIN

27.4%; Score 550.6; DB 28; Length 784;

Query Match Best Local Similarity 83.8%; Pred. No. 1.7e-99;
Matches 651; Conservative 0; Mismatches 114; Indels 12; Gaps 2;

QY 123 TCTCTCTTCTCTCACTTACTTGTGATCCGCCATGAGCTACGAACCTGAAAGCT 182
Db 774 TCTCTCTTCTCTCACTTACTTGTGATCCGCCATGAGCTACGAACCTGAAAGCT 715
QY 183 CCTTAATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 242
Db 714 CCTTAATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 655
QY 243 CCTTCTGCGCGGCAATCATCTGCGGCAATCATCTGCGGCAATCATCTGCGGCAATCATCT 302
Db 654 CTAATCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 595
QY 303 GTGACCTGCTTGGGGAATACCTTCTGCGGCTTAAATATCCGTTATACAAACCG 362
Db 594 GTGACCTGCTTGGGGAATACCTTCTGCGGCTTAAATATCCGTTATACAAACCG 544
QY 363 GAAAGCTTATCCCGGTCGAAAGAGCTTCTTGTGTTGGAAGATGATCATGATGCA 422
Db 544 GAAAGCTTATCCCGGTCGAAAGAGCTTCTTGTGTTGGAAGATGATCATGATGCA 484
QY 484 GCGCTTACGACCAACCGAATGCGAAGAGAGCGGAGATGCGAAGAGATGCGAAGAGATGCGA 424
Db 423 GCACTGATGCTACCGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 482
QY 483 TGGCTTTGAGCTTGAAGAGAGCTCGGCTGATGATGATGATGATGATGATGATGATGATGAT 542
Db 423 TGGCTTTGAGCTTGAAGAGAGCTCGGCTGATGATGATGATGATGATGATGATGATGATGAT 364
QY 543 AGATTGATTAAGACCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 602
Db 363 AGATTGATTAAGACCGGCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 304
QY 603 TGTTTAAGAGCAATTTGTTTGAACCAAGGTTTAACTGGGAGAGCTTGGCGGTA 662
Db 303 TGTTTAAGAGCAATTTGTTTGAACCAAGGTTTAACTGGGAGAGCTTGGCGGTA 244

QY 663 TCCTTCGAAACATCTCTTGTAGTACAAACAAATCGAAGAGCCGAGACGCAACGAG 722
 DB 243 TCCTTCGAAACATCTCTTGTAGTACAAACAAATCGAAGAGCCGAGACGCAACGAG 184
 QY 723 TGAATCAGACGAGATGCTGAGTCTTCTGAAACAG---AGTACTACGAAACCTGTT 779
 DB 183 TGAATCAGACGAGATGCTGAGTCTTCTGAAACAG---AGTACTACGAAACCTGTT 124
 QY 780 TTGTCGTGAGTGTCTTAAACGCGCTGCTTACAAACATGATGCTCTGTATTTGAGC 839
 DB 123 TTGTCGTGAGTGTCTTAAACGCGCTGCTTACAAACATGATGCTCTGTATTTGAGC 64
 QY 840 AAGAGTATGAGCTTGAACAAACATGATGCTTACAAACATGATGCTCTGTATTTGAGC 896
 DB 63 AAGAGTATGAGCTTGAACAAACATGATGCTTACAAACATGATGCTCTGTATTTGAGC 7

 RESULT 3
 B2428529 809 bp DNA linear GSS 13-DEC-2002
 LOCUS B2428529 BO_1.6_2_KB_tot Brassica oleracea genomic clone BONHW74,
 DEFINITION genomic survey sequence.
 ACCESSION B2428529
 VERSION B2428529.1 GI:26671510
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids
 ; eucosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 809)
 Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 UNPUBLISHED
 OTHER_GSS: BONHW74TF
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..809
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BONHW74"
 /clone_1lb="BO_1.6_2_KB_tot"
 /note="Vector: pHO51; Site 1: BstXI; 1.6-2 kb sheared
 total DNA inserted into pHO51 using BstXI linkers"
 BASE COUNT 241 a 216 c 146 g 206 t
 ORIGIN
 Query Match 25.0%; Score 502.6; DB 29; Length 809;
 Best Local Similarity 86.8%; Pred.No. 6.3e-90;
 Matches 553; Conservative 0; Mismatches 84; Indels 0; Gaps 0;
 QY 1342 TTGAGTATTATTACAGAGATGATTTTCAGAGAAACAGACAGCTGGCTCTATCG 1401
 DB 758 TTATATTATGTCAAGAGATGATTTTCAGAGAAACAGACAGCTGGCTCTATTCG 699
 QY 1402 AGTGATCTCGTACAGATGCTCTTCAATCCAGATATGCAATCAACGATCAAAAACGAGC 1461
 DB 698 AGTGAACCTGATGATGATGCTCTTCAATCCAGATATGCAATCAACGATCAAAAACGAGC 639
 QY 1462 TGAATCAAGTATGCTGGGAAATCAAGAGCCCTTATGATGATGCTGCTTCACTTCAT 1521
 DB 638 TTGATTTGGTGTGGAAATCAAGAGCTGATGATGATGCTGCTTCACTTCAT 579

QY 1522 ATCTACGGCTGTGTGTAAGAGATATTAGGCTTATCTCTCCAGGCCCCACTTCTATCAT 1581
 DB 578 ATCTACGGCTGTGTGTAAGAGATATTAGGCTTATCTCTCCAGGCCCCACTTCTATCAT 519
 QY 1582 GGGCCCGCTTGGCCATTAACAGACAGATGCTTATGCTGCTTCTGCTCCGACAGGACCA 1641
 DB 518 GGGCCCGCTTGGCCATTAACAGATGCTTATGCTGCTTCTGCTCCGACAGGACCA 459
 QY 1642 CAGCAATGATGAACATGTGGCCATATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 1701
 DB 458 CAGCAATGATGAACATGTGGCCATATGCTGCAATGCTGCAATGCTGCAATGCTGCAAT 399
 QY 1702 AGTTAAACCTGAGAGGTTGCTGCAAAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 1761
 DB 398 AGTTAAACCTGAGAGGTTGCTGCAAAAGAGAGTGAAGTGAAGTGAAGTGAAGTGAAGT 339
 QY 1762 CGAATTTGAGACTTGACCTTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 1821
 DB 338 CGAATTTGAGACTTGACCTTTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTGCGGCTG 279
 QY 1822 GTTTTACTACCGCTTATGTTTGAACGCGATGATGATGATGATGATGATGATGATGAT 1881
 DB 278 GATTGACAAACCTTACGTTGACCGCGGCTGTTACAGAAATTTGAATGAGGAGTGT 219
 QY 1882 CCGATGTACAGCGGCTTACTTATCTGAGAACTGAGGCTTCTTGTGCGAGATGCTATATC 1941
 DB 218 CAGATGTACAGCGGCTTACTTATCTGAGAACTGAGGCTTCTTGTGCGAGATGCTATATC 159
 QY 1942 CTCTTCTCTTAAATTTGCGGCTGAGCGGCTTAAAT 1978
 DB 158 CTCTTCTCTTAAATTTGCGGCTGAGCGGCTTAAAT 122

 RESULT 4
 BH451907/c
 LOCUS BH451907 BOGS044TR BOGS Brassica oleracea genomic clone BOGS044, genomic
 DEFINITION survey sequence.
 ACCESSION BH451907
 VERSION BH451907.1 GI:17637618
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons; rosids
 ; eucosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 674)
 Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 UNPUBLISHED
 OTHER_GSS: BOGS044TF
 CONTACT: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..674
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOGS044"
 /clone_1lb="BOGS"
 /note="Vector: pHO51; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHO51 using BstXI linkers"
 BASE COUNT 191 a 194 c 130 g 155 t
 ORIGIN
 FEATURES
 source

Query Match 23.8%; Score 479; DB 28; Length 674;
 Best Local Similarity 89.6%; Pred. No. 3.2e-85;
 Matches 515; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

1405 GGATCTCGTAGAGATGCTCTTCATCAGATATGACATACAGGATCAAAACGAGCTGG 1464
 674 GGATCTCGTAGAGATGCTCTTCATCAGATATGACATACAGGATCAAAACGAGCTGG 615

1465 ATCAAGTAGTGGGAATCAAGACCTTAAGTGAATGATCTTGGCTTCACTTCAATTC 1524
 614 ATTCGATGTCGGAATAACAGACCGCTGATGATGATGATGATGATGATGATGATG 555

1525 TAACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1584
 554 TGACGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 495

1585 CCCGTTGGCCATTAACAGACGATCTTGTATGATGATGATGATGATGATGATGATG 1644
 494 CCCGTTGGCCATTAACAGACGATCTTGTATGATGATGATGATGATGATGATGATG 435

1645 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1704
 434 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 375

1705 TTAACCTGAGAGGTTCTGTCGAAAGAGGTTGAGTGAAGTTCCTGTTCTTGGGTCGG 1764
 374 TTAACCTGAGAGGTTCTGTCGAAAGAGGTTGAGTGAAGTTCCTGTTCTTGGGTCGG 315

1765 ATTGAGACTGACACTTCCGATGATGATGATGATGATGATGATGATGATGATGATG 1824
 314 ACTTGAAGCTGAGGCTCTTGGGTCGATGATGATGATGATGATGATGATGATGATGATG 255

1825 TTACTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1884
 254 TGACTACCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 195

1885 ATGTAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1944
 194 TTGGAACGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 135

1945 TTCTGCTAAATGCGCCGTAGCCGAGTTAAAA 1979
 134 TTGCTGCTAAATGCGCCGTAGCCGAGTTAAAA 100

RESULT 5
 BH431683 804 bp DNA linear GSS 12-DEC-2001
 LOCUS BOHRT46R BOHR Brassica oleracea genomic clone BOHRT46, genomic
 DEFINITION survey sequence.
 ACCESSION BH431683
 VERSION BH431683.1 GI:17617404
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 1 (bases 1 to 804)
 1 eucoside II; Brassicales; Brassicaceae; Brassica.
 REFERENCE Town, C.D., Van Aken, S., Utecherback, T., Koo, H. and Fraser, C.M.
 TITLE Whole genome shotgun sequencing of Brassica oleracea
 JOURNAL Unpublished
 COMMENT Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..804

Query Match 23.5%; Score 472; DB 28; Length 804;
 Best Local Similarity 77.1%; Pred. No. 8.1e-84;
 Matches 628; Conservative 0; Mismatches 150; Indels 36; Gaps 3;

761 AGTAGTAAGAACCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATG 820
 778 AGTAGTAAGAACCTGTTTGTGATGATGATGATGATGATGATGATGATGATGATG 719

821 ATGTGCTCTGATTTGGGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 880
 718 ATGTGCTCTGATTTGGGCAAGAGATGATGATGATGATGATGATGATGATGATGATG 660

881 ATGTGCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 940
 659 TTGTCGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 600

941 CTATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1000
 599 TTATCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 540

1001 ACCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1060
 539 ACCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 480

1061 GATTGCTGAGGTTTGTGCTCTGCTCCATGATGATGATGATGATGATGATGATGATG 1120
 479 GATTGCTGAGGTTTGTGCTCTGCTCCATGATGATGATGATGATGATGATGATGATG 420

1121 ATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1180
 419 ATGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 385

1181 TGAATAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1240
 384 TGAATAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 325

1241 TGTGATTAATTAACATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 1300
 324 GATGATTAATTAACATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 265

1301 TTTATAACAGTTTCTATTAATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 1360
 264 TTTATAACAGTTTCTATTAATAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAA 215

1361 GATGATTAATTAACATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 1420
 214 AATGATTAATTAACATTAAGATGATGATGATGATGATGATGATGATGATGATGATG 155

1421 GATCCTTCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1480
 154 GATCCTTCATCAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 95

1481 ATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1540
 94 ATCAAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 35

1541 AGAGTATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1574
 34 AGAGTATGAGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1

RESULT 6

BH510149/c 745 bp DNA linear GSS 13-DEC-2001
 LOCUS BOGT049TR BOGT Brassica oleracea genomic clone BOGT049, genomic
 DEFINITION survey sequence.
 ACCESSION BH510149
 VERSION BH510149.1 GI:11718239
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eucosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 745)
 REFERENCE Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
 AUTHORS Whole genome shotgun sequencing of Brassica oleracea
 TITLE Unpublished
 JOURNAL Other GSSs: BOGT049TF
 COMMENT Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..745
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOGT049"
 /clone_1lb="BOGT"
 /note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOSt using BstXI linkers"
 BASE COUNT 219 a 186 c 137 g 203 t
 ORIGIN
 Query Match 22.8%; Score 458.8; DB 28; Length 745;
 Best local similarity 78.2%; Pred. No. 3.5e-81;
 Matches 568; Conservative 0; Mismatches 147; Indels 11; Gaps 1;
 Class: sheared ends.
 Location/Qualifiers
 1..745
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOGT049"
 /clone_1lb="BOGT"
 /note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOSt using BstXI linkers"
 BASE COUNT 219 a 186 c 137 g 203 t
 ORIGIN

DB 306 CCTCAGGCCCACTTCTGTCGTGGGCCCGCTTGGCCATACAGACAGATCGTTGATGTT 247
 QY 1679 CGTTTGTTCGGAGAGGACACAGACGATGTAACATGAGGCGGATCGATGATCA
 DB 246 TGTCTTGTTCGGGCGGAGACACAGACGATGTAACATGAGGCGGATCGATGATCA 167
 QY 1739 CACGTGTGGGTGATCTCTTGAAGTTTAACTGAGAGGTTCTGGCAAAAGAGTGAG
 DB 127 CACGTGTGGGTGATCTCTTGAAGTTTAACTGAGAGGTTCTGGCAAAAGAGTGAG
 QY 1799 GTGAGATTGTTGGTTCTTGGGTCGAGATTGAGATTGACCTTGGGTCGGTCGCG
 DB 67 GTGAGATTGTTGGTTCTTGGGTCGAGATTGAGATTGACCTTGGGTCGGTCGCG
 QY 1859 ATTGCCCCGGAGAAATCTTGTGTTTACTACCTTATGTTTGGACGGCGATGATTTA
 DB 7 ATGTCCTGGGAGAAATCTTGTGTTTACTACCTTATGTTTGGACGGCGATGATTTA
 QY 1865 CATGAG 1865
 DB 6 CATGAG 1
 RESULT 7
 BH527239/c 727 bp DNA linear GSS 13-DEC-2001
 LOCUS BOHB06TR BOHB Brassica oleracea genomic clone BOHB06, genomic
 DEFINITION survey sequence.
 ACCESSION BH527239
 VERSION BH527239.1 GI:17735324
 KEYWORDS GSS.
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 ; eucosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 727)
 REFERENCE Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
 AUTHORS Whole genome shotgun sequencing of Brassica oleracea
 TITLE Unpublished
 JOURNAL Other GSSs: BOHB06TF
 COMMENT Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TR
 Class: sheared ends.
 Location/Qualifiers
 1..727
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOHB06"
 /clone_1lb="BOHB"
 /note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOSt using BstXI linkers"
 BASE COUNT 197 a 204 c 145 g 181 t
 ORIGIN
 Query Match 21.0%; Score 421.6; DB 28; Length 727;
 Best local similarity 80.8%; Pred. No. 9.1e-74;
 Matches 506; Conservative 0; Mismatches 114; Indels 6; Gaps 1;
 Class: sheared ends.
 Location/Qualifiers
 1..727
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO100DH3"
 /db_xref="taxon:3712"
 /clone="BOHB06"
 /clone_1lb="BOHB"
 /note="Vector: pHOSt; Site 1: BstXI; 2-3 kb sheared
 genomic DNA inserted into pHOSt using BstXI linkers"
 BASE COUNT 197 a 204 c 145 g 181 t
 ORIGIN

Tue Jan 20 17:43:19 2004

us-10-022-025a-1.rst

Page 6

583 GATGTCCTTCATCAAGATCTAGACACACAGTTTCAACAGAGCTTGACAGGTCGTAG 524
1478 GAAATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCTATCTAAGCGCTGTGT 1537
523 GAGATCAAGAGCCCTAGATGAATCTGACTTGGCTTCACTTCTATCTAAGCGCTGTGT 464
1538 GAAAGAGTATTGAGGCTTCACTTCCAGGCCCACTTCTATCTAAGCGCTGTGT 1597
463 CAAAGAGTCTTGAGGCTTCACTTCCAGGCCCACTTCTATCTAAGCGCTGTGT 404
1598 AACAGACAGATCGTTAGTGTGTCTTTGTCGCGCAGGACCAACAGCAATGTGTAACT 1657
403 CACAGATACGATCAATTGACGCGCGTGTGTGCGCGAGGACCACTGCAATGTGTAACT 344
1568 GTGGCCCGATCGCATGATGTCACACGCTGTGGGTTGATCTTGTGAGTTTAACTGAG 1717
343 GTGGGCTATGAGCGATGATGTCACACGCTGTGGGTTGATCTTGTGAGTTTAACTGAG 284
1718 GTTGTGAGCAAAAGAGTGTGAGTGTGCTTGTGCTTGTGCTGAGTTTGAAGCTTGC 1777
283 GTTGTGAGCAAAAGAGTGTGAGTGTGCTTGTGCTTGTGCTGAGTTTGAAGCTTGC 224
1778 ACCCTTGGGCTGCGGCTGTGCTGAGTTTGTGCGCGAGCACTTGTGCTGAGTTT 1837
223 ACCCTTGGGCTGCGGCTGTGCTGAGTTTGTGCGCGAGCACTTGTGCTGAGTTT 164
1838 GTTGTGAGCAAAAGAGTGTGAGTGTGCTTGTGCTTGTGCTGAGTTTGAAGCTTGC 1891
163 GTTGTGAGTGTGAGTGTGAGTGTGCTTGTGCTTGTGCTGAGTTTGAAGCTTGC 104
1892 CGGCTTACCTTATCTGAGAACTGAGGCTTGTGCTGAGTGTGCTTGTGCTGAGTTT 1951
103 GACCTTACCTTATCTGAGAACTGAGGCTTGTGCTGAGTGTGCTTGTGCTGAGTTT 44
1952 TAAATTGCGCGCTGAGGCGCAGTTAA 1977
43 TAAATTGCGCGCTGAGGCGCAGTTAA 18

RESULT 8 659 bp DNA linear GSS 13-DEC-2002
B246758/c
LOCUS BONND11R BO.1.6.2 KB tot Brassica oleracea genomic clone BONND11,
DEFINITION genomic survey sequence.
ACCESSION B246758
VERSION B246758.1 GI:26759933
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; euroside II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 659)
Towm,C.D., Van Aken,S., Uterback,T., Koo,H. and Frazer,C.M.
AUTHORS
TITLE Whole genome shotgun sequencing of Brassica oleracea
JOURNAL
COMMENT Other GSSs: BONND11TF
Contact: Chris Town
TIGR
5712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES
Source
1.659
Location/Qualifiers
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BONND11"

/clone_1id="BO.1.6.2_KB_tot"
/note="Vector: PHOS1; Site: 1; BstXI, 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 172 a 143 c 195 g 149 t
ORIGIN
Query Match 20.9%; Score 419.6; DB: 29; Length 659;
Best Local Similarity 82.2%; Pred: No. 2.3e-73;
Matches 549; Conservative 0; Mismatches 99; Indels 20; Gaps 5;

78 TCATCTTAAAGCCCAAGAGCC-AAAAGACACTCTCTTTGTT-----TATCTCTC 130
659 TCATCTTAAAGCCCAAGAGCCAAAGACACTCTCTTTGTTCTCTCTCTCTCT 600
131 TTCTCTCACTTACTTCTCTTCTGATCCGCCAT--GGCTACAAACTGAAAGC--TCCTT 186
599 ATAGCTTCACTTCACTTCTCTTCTGATCCGCCATCCCATGAGGCTTACAACTGAAAGCTTCCCT 540
187 ATCTTGGCCCTTGTGCAATGAGAGCGCTTCAAGCCAAACCAACTTGCCTTCTCCCT 246
539 GATCTTGGCCCTTGTGCAATGAGAGCGCTTCAAGCCAAACCAACTTGCCTTCTCTAT 480
247 CCTGCGCGTCAACATCTGAGCTGCGCATATCTCTCTTCTTATGAACTTATCCCGGTG 306
479 ACTTGCATCTGATATGCTGCTGCTGCTTCTCTCTCTCTCTGAGCTTATCCAGGTG 420
307 ACCCTGTTGGGGGAAATACCTCTTGGCGGTTATATCCGTTTCAACAAACCGGAAA 366
419 ACCCTGATGAGGCAATACCTCTTGGCGGTTATATCCGTTTCAACAAACCGGAAA 369
367 CGTTATCCCGTCCAAAGAGGCTTCTTGGTTGGAAGCATGTCACTCATGTCAAGAC 426
368 GGTATTCCTCGGTTCCAAAGAGGCTTCTTGGTTGGAAGCATGTCTCATGTCCAGGCG 309
427 TCTAGCTCAACCAAGATGTGATGCAAGCTTGAAGAAATTCGAGGCAAGAGGCTATGCG 486
308 TCTAGCAACCAACCAAGATGTGATGCAAGCTTGAAGAAATTCGAGGCAAGAGGCTATGCG 249
487 TTTCAGCTTAAAGAGAGACTCGCGTATCTCTCACTGCAATCCCAAGTATGCGAAAGAT 546
248 TTTCAGCTTAAAGAGAGAGACTCGCGTATCTCTCACTGCAATCCCAAGTATGCGAAAGAT 189
547 TCTGAATAGCCGCTTTTGTGATGACCGGTTAAAGATGCGGCTTCTCACTGATGT 606
188 TCTGAATAGCCGCTTTTGTGATGACCGGTTAAAGATGCGGCTTCTCACTGATGT 129
607 TAAACAGCAATGTGTTTGCACACACAGGTTTACTGCGCAACGCTTGCCTGATGCG 666
128 TAAACAGCAATGTGTTTGCACACACAGGTTTACTGCGCAACGCTTGCCTGATGCG 69
667 TTGAACCACTCTTCTTAAAGCAATCAAGAGCCGAGAGCCCAACAGAGCTGAT 726
68 TTGAACCACTCTTCTTAAAGCAATCAAGAGCCGAGAGCCCAACAGAGCTGAT 9
727 CTCAAGCC 734
8 CGCTAGCC 1

RESULT 9 776 bp DNA linear GSS 13-DEC-2002
B2443616
LOCUS BONF28TR BO.1.6.2 KB tot Brassica oleracea genomic clone BONF28,
DEFINITION genomic survey sequence.
ACCESSION B2443616
VERSION B2443616.1 GI:26703871
KEYWORDS
SOURCE
ORGANISM
Brassica oleracea
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
; euroside II; Brassicales; Brassicaceae; Brassica.
1 (bases 1 to 776)

Tue Jan 20 17:43:19 2004

us-10-022-025a-1.rst

Page 7

AUTHORS Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
TITLE Whole genome shotgun sequencing of *Brassica oleracea*
JOURNAL Unpublished
COMMENT Other_GSSs: BONFP28TR
Contact: Chris Town

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES
source location/Qualifiers

1. 776
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="T0100DH3"
/db_xref="taxon:3712"
/clone="BONFP28"
/clone_1lb="BO.1.6.2 KB tot"
/note="Vector: PHOS1, Site 1: BstXI; 1.6-2 kb sheared
total DNA inserted into PHOS1 using BstXI linkers"
BASE COUNT 196 a 156 c 207 g 217 t
ORIGIN

Query Match 19.7%; Score 396.6; DB 29; Length 776;

Best Local Similarity 74.2%; Pred. No. 8.9e-69;

Matches 529; Conservative 0; Mismatches 179; Indels 5; Gaps 2;

QY 1230 ATGTACGCTGTGTGAATTAATTAAGCACTTGAAGTTTGTGACATCAATTTGCT 1289
DB 64 ACGTTCCTATTATTTATTTTAAATTAACAAACAAACATATATATCAATTTTAA 123
QY 1290 AATTTTATGATTTTATTAACAGTTCTATATAAAGTAAAGTAAATTTTGAATTA 1349
DB 124 ATATATCCATTTTAGTATATGACACAGCTTATATGATTAACGTAATTCATTTAG 183
QY 1350 TTATTAACAGAGATGATATTCAGAGACAGACAGATTCGCTTAAATGAGTGAATC 1409
DB 184 GTTTTACAGAAATGATATTTAGAGACAGACAGCTGCTGATTTGATCGAGTGAAT 243
QY 1410 CTGCTAGAGATGCTCTTCAATCCAGATGCAATCAAGGTCAAAAGAGCTGATCAA 1469
DB 244 CTGCTAGAGATGCTCTTCAATCCAGATGCAATCAAGGTCAAAAGAGCTGATCAA 303
QY 1470 GTAGTCGGAATCAAGAGCCCTAGATGATGATCTGATCTGCTTCAATCTTAAAG 1529
DB 304 GTGTAAGGAGATCAAGAGCCGCGAAGATCTGAGCTGCTTCAATCTTAAAGCTG 363
QY 1530 GCTGTGATGAAGAATGATTTAGGCTTCACTCCAGGCGCACTTTATCAATGGGCGCT 1589
DB 364 GCTGTGATGAAGAATGATTTAGGCTTCACTCCAGGCGCACTTTGCTGAGGCGCT 423
QY 1590 TTGGCCATTAACAGACAGATGCTGATGCTGCTTCTGCTCCGAGGAGCAACAGCATG 1649
DB 424 TTGCAATCAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 483
QY 1650 GTAAACATGTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 1709
DB 484 GTTAAACATGTGGCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 543
QY 1710 CCGAGAGGTTGCGGCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATG 1769
DB 544 CCGAGAGGTTGCGGCAAAAGAGGATGATGATGATGATGATGATGATGATGATGATG 603
QY 1770 AGACTTGACCTTTCCGCTGCGGTCTGCGATTTGCGCGGAAAGATCTTGATTTACT 1829
DB 604 AGACTTGACCTTTCCGCTGCGGTCTGCGATTTGCGCGGAAAGATCTTGATTTACT 663
QY 1830 ACGTTATGTTTGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 1886
DB 664 ACGTTATGTTTGAAGGAGGATGATGATGATGATGATGATGATGATGATGATGATG 723

QY 1887 GGTAAGG--CGTTGACTTATGAGAACTGAGGCTTTCTGCGAGATGCT 1937
DB 724 GATGACAGACCTTTGACTTCTCGAGAACTGAGGCTCTGCTGATGATGCT 776

RESULT 10

BR451899

LOCUS

DEFINITION

BOG5044TR BOGS Brassica oleracea genomic clone BOG5044, genomic

survey sequence.

ACCESSION

BR451899

VERSION

BR451899.1

KEYWORDS

GSS.

ORGANISM

Brassica oleracea

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids

; eurosids II; Brassicales; Brassicaceae; Brassica.

1 (bases 1 to 576)

Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.

Whole genome shotgun sequencing of *Brassica oleracea*

Unpublished

Other_GSSs: BOG5044TR

Contact: Chris Town

TIGR

9712 Medical Center Drive, Rockville, MD 20850, USA.

Tel: 301-838-3523

Fax: 301-838-0208

Email: cdtown@tigr.org

DNA is from a doubled haploid provided by Tom Osborn.

Seq primer: TR

Class: sheared ends.

FEATURES

source location/Qualifiers

1. 576

/organism="Brassica oleracea"

/mol_type="genomic DNA"

/strain="T0100DH3"

/db_xref="taxon:3712"

/clone="BOG5044"

/note="Vector: PHOS1, Site 1: BstXI; 2-3 kb sheared

genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 137 a 162 c 114 g 163 t

ORIGIN

Query Match 19.0%; Score 382.6; DB 28; Length 576;

Best Local Similarity 83.1%; Pred. No. 5.5e-66;

Matches 466; Conservative 0; Mismatches 79; Indels 16; Gaps 2;

QY 29 TGTATTAATGATGCTGATTAATCTATATAATACCAAGCATCTGCTATCTTTAG 88
DB 25 TTTATTTTAAAGATTCGATTTTCTATATTAAGCAAGATCTGTTGATCTTTAG 84
QY 89 CCCCAGAGCCCAAAAGCACTCTCTTTGTTTATCTCTCTCTCTCTCTCTCTCTCTCT 141
DB 85 CCCCAGAGCCCAAAAGCACTCTCTTTAGTTCTCTCTCTCTCTCTCTCTCTCTCT 144
QY 142 TACTTTGTTTGAATCCGATGCTAGAAAGCTGAAAGCTCCCTATCTTTGCGCTTTT 201
DB 145 TGCATCTCTTTGATCCCATGCTGATGATGATGATGATGATGATGATGATGATG 204
QY 202 GTCCAAATGACAGCTTTAAGCCCAACCAACTTCTCTCTCTCTCTCTCTCTCTCTCA 261
DB 205 GTCCAAATGACAGCTTTAAGCCCAACCAACTTCTCTCTCTCTCTCTCTCTCTCTCA 264
QY 262 CATCTGCTGCGCATATCTCTCTCTTATGAGCACTATCCGCGTGAAGCTTTGCGGGA 321
DB 265 GCTGTGCTGCGCTTTCTCTCTCTTCTGAGCTTATCAAGTGAAGCTGATGAGGGA 324
QY 322 ATACTCTTCCGCGGTTAATATCCGTTTACAAACCGGAAACGTTATTCGCGTCC 381
DB 325 ATACTCTTCCGCGGTTAATATCCGTTTACAAACCGGAAACGTTATTCGCGTCC 375

QY	Db	QY	Db	RESULT 11
382	AAAAAGCTTCCTCCCTTGGTTGGAAGCATGTCACTCATGTCAAGCACTTACTGACGAGC	441		
376	AAGAGAGTTCCCTTCCTTGGGAGACATGTCTCTCATGTCCAGCGCTTACGACACACCG	435		
442	AATGCTGATGCAGCTGAGAAATTCGAGGCCAGAGGCTCATGTGCTTTCACTTAGAGA	501		
436	AATGCGAAGACACCCGAGATCGAGAGCCAAACGGCTCATGTGCTTTAGCTTAGGCGA	495		
502	GACTCGGGTATTCGTCACGTGCAATCCCGACGTGCGGAAAGAGATTCGAAATAGCCGGT	561		
496	GACACGGGTATTCGTACGTGCAACCCGACGTGCTTAAGAGATTCTGAATAGTCCAGT	555		
562	TTTTCGTATGCACCGGTTAA	582		
556	TTTTCGTACCGCTCGGTTAA	576		
BH977803	BH977803	705 bp	DNA	linear
LOCUS	ode14a11.b1			
DEFINITION	Brassica oleracea002			
ACCESSION	BH977803			
VERSION	BH977803.1			
KEYWORDS	GSS.			
SOURCE	Brassica oleracea			
ORGANISM	Brassica oleracea			
REFERENCE	Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids			
TITLE	; eusoids II; Brassicales; Brassicaceae; Brassica.			
JOURNAL	1 (bases 1 to 705)			
COMMENT	DeLahunty, K., Fewell, G., Fulcon, L., McCombie, W.R., Miner, T., Nash			
	W., Rahnowicz, P.D. and Wilson, R.K.			
	Whole genome shotgun reads from Brassica oleracea			
	Unpublished			
	Contact: Richard K. Wilson			
	Genome Sequencing Center			
	Washington University School of Medicine			
	Email: submissions@watson.wustl.edu			
	Plate: ode14 row: a column: 11			
	Seq primer: -21upor forward			
	Class: shotgun			
	High quality sequence start: 18			
	High quality sequence stop: 519.			

FEATURES		Source	Location/Qualifiers
			1..705
			"/organism="Brassica oleracea"
			"/mol_type="Genomic DNA"
			"/db_xref="taxon:3712"
			"/clone_id="B.oleracea002"
			"/note="Vector: pOTW13: Whole genome shotgun library from flowering buds. DNA was purified from a crude nuclear prep using Brasica oleracea T01000DH3 buds, provided by Thomas Osborn at the University of Wisconsin. Genomic DNA was provided by Pablo Rabinowicz (CSH) and the shotgun library prepared at Washington University Genome Sequencing Center."
BASE COUNT	173 a	143 c	193 g 193 t 3 others
ORIGIN			
Query Match		16.8%;	Score 337.6; DB 28; Length 705;
Best Local Similarity		76.1%;	Pred. No.5.2e-57;
Matches 466; Conservative		0;	Mismatches 141; Indels 5; Gaps 4;
QY	1340	TTTTGAGTATTATTACAGAGATGATATTTCGAGGACACACACAGTTGCGGCTTTAAT	1399
DB	93	TTTTTTGTGGCTTTTACAGGAATGATATTTCGAGGAATGACACGGGTGGCTGTTGAT	152
QY	1400	CGAGTGCATCTCGCTAGGATGGTCTCTTATCATCAAGTATGCATCAACGGTACAAACGA	1455
DB	153	CGAATGCAATTCGTGATATGATGGCTCTTATCATCAAGTATTCGTAATGGTTCAACACGA	212

QY	1460	CTCGATGATCAAGTAGTGTGGGGAAATCAAGAGCCCTAGATGATCTGAAGCTTGACCTCC	1519
Db	213	GCTTGATCAGAACGTGGGAAGATCAGACGCCCTGGAAGATCTGAAGTGGCGCTCTTAC	272
QY	1520	ATATCTAACGGCTGTGGTGAAAGAAATTTGAGGCT-TCATCTCCAGGCCAC-TTCTA	1577
Db	273	ATATCTACAGCTGTGTGAAAGAAAGCTTGAAGCTTTCACCCGACAGGCCACTTCTG	332
QY	1578	TCATGGGCGCGTTTGGCCATAAGACAGCATGTTGATGTGCTCTTGTTCGGCAGGG	1637
Db	333	TGTGGGCGCGCTTAGCAATCAAGACACAAATATTGACGGCGGTCTGTACCGGAGGT	392
QY	1638	ACCAACAGCAATGTGTGAACATGTGGGCGGATTCGATATCAACAGTGTGGGTTGATCTT	1697
Db	393	ACCAACCAATGTGTGAACATGTGGGCTATATAGCACTATATCAACAAAGTGTGGAAATCTT	452
QY	1698	TTGGAAGTTTAAACCTGAGAGTTCGTGGCAAAAAGAGTGAGGAGTTTCGGTCTT	1757
Db	453	TTGGAAGTTTAAACCTGACCGTTTGTACAAAGAAATGAGAGGGGAGTTCTCGGTTCTT	512
QY	1758	GGGTCGAGTTTGACACTTGACCTTTTCGGGTGGGTCTCGGATTTGCCCCGGAGAAT	1817
Db	513	GGATCGGATCTGAAGGCTGACCTTTCGGGGCGGGCCGTGGGGCTCCCTCGAAAGAATT	572
QY	1818	CTTGCTTTTACTACCGTTATGTTTGTGAACGGCGAGTAGATGTACATGATTTGAATGG--G	1875
Db	573	CTAAGGGTTAGCACCCTGATGTGTATTTGGATCGCACTCTCTTGCCCGAATNAGATGGTGG	632
QY	1876	GACCGTCGATGTGTAAACGGCGTTGACATTATCTGAGAA-CTGAGCGTTTCTTGGAGATG	1934
Db	633	CACCTAACGGGAAAAGACTGTGACTTTGTCGAGAAATGAGGCTCTTCCCTGTGAGATG	692
QY	1935	GCTATATCTCTT 1946	
Db	693	GCTATATCTCTT 704	

RESULT 12
 BH534160
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 732 bp DNA linear GSS 14-DEC-2001
 survey sequence.
 BOHC239TF BOHC Brassica oleracea genomic clone BOHC239, genomic
 BH534160
 BH534160.1 GI:17761104
 GSS.
 Brassica oleracea
 Brassica oleracea
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 i; eurosid II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 732)
 Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Fraser,C.M.
 Whole genome shotgun sequencing of Brassica oleracea
 Unpublished
 Other GSSs: BOHC239TF
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seq primer: TF
 Class: sheared ends.
 Location/Qualifiers
 1..732
 /organism="Brassica oleracea"
 /mol_type="genomic DNA"
 /strain="TO1000DH3"
 /db_xref="taxon:3712"
 /clone="BOHC239"
 /clone_1ib="BOHC"
 /note="vector: pHOSt, site_1: BstXI; 2-3 kb sheared

Db 576 ACCGGGTGATCTATCGAAGTGTGAGCTATCTCGAATGACACCCCTGACTG 635
 QY 1951 CTAAATGGCGCGGTAGCGGCACTTAAA 1978
 Db 636 TGAAGGTGGCGCCGAGGCAACATGATA 663

RESULT 14
 CB343993 638 bp mRNA linear EST 10-APR-2003
 LOCUS CA48N0001.11BF G02 Cabernet Sauvignon leaf-CA48N vltis vinifera
 DEFINITION cDNA clone CA48N0001.11BF_G02 5', mRNA sequence.

ACCESSION CB343993
 VERSION CB343993.1 GI:28964960
 KEYWORDS EST
 SOURCE vltis vinifera
 ORGANISM Vitis vinifera

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 2 Vitis vinifera
 3 Vitis vinifera
 4 Vitis vinifera
 5 Vitis vinifera
 6 Vitis vinifera
 7 Vitis vinifera
 8 Vitis vinifera
 9 Vitis vinifera
 10 Vitis vinifera
 11 Vitis vinifera
 12 Vitis vinifera
 13 Vitis vinifera
 14 Vitis vinifera
 15 Vitis vinifera
 16 Vitis vinifera
 17 Vitis vinifera
 18 Vitis vinifera
 19 Vitis vinifera
 20 Vitis vinifera
 21 Vitis vinifera
 22 Vitis vinifera
 23 Vitis vinifera
 24 Vitis vinifera
 25 Vitis vinifera
 26 Vitis vinifera
 27 Vitis vinifera
 28 Vitis vinifera
 29 Vitis vinifera
 30 Vitis vinifera
 31 Vitis vinifera
 32 Vitis vinifera
 33 Vitis vinifera
 34 Vitis vinifera
 35 Vitis vinifera
 36 Vitis vinifera
 37 Vitis vinifera
 38 Vitis vinifera
 39 Vitis vinifera
 40 Vitis vinifera
 41 Vitis vinifera
 42 Vitis vinifera
 43 Vitis vinifera
 44 Vitis vinifera
 45 Vitis vinifera
 46 Vitis vinifera
 47 Vitis vinifera
 48 Vitis vinifera
 49 Vitis vinifera
 50 Vitis vinifera
 51 Vitis vinifera
 52 Vitis vinifera
 53 Vitis vinifera
 54 Vitis vinifera
 55 Vitis vinifera
 56 Vitis vinifera
 57 Vitis vinifera
 58 Vitis vinifera
 59 Vitis vinifera
 60 Vitis vinifera
 61 Vitis vinifera
 62 Vitis vinifera
 63 Vitis vinifera
 64 Vitis vinifera
 65 Vitis vinifera
 66 Vitis vinifera
 67 Vitis vinifera
 68 Vitis vinifera
 69 Vitis vinifera
 70 Vitis vinifera
 71 Vitis vinifera
 72 Vitis vinifera
 73 Vitis vinifera
 74 Vitis vinifera
 75 Vitis vinifera
 76 Vitis vinifera
 77 Vitis vinifera
 78 Vitis vinifera
 79 Vitis vinifera
 80 Vitis vinifera
 81 Vitis vinifera
 82 Vitis vinifera
 83 Vitis vinifera
 84 Vitis vinifera
 85 Vitis vinifera
 86 Vitis vinifera
 87 Vitis vinifera
 88 Vitis vinifera
 89 Vitis vinifera
 90 Vitis vinifera
 91 Vitis vinifera
 92 Vitis vinifera
 93 Vitis vinifera
 94 Vitis vinifera
 95 Vitis vinifera
 96 Vitis vinifera
 97 Vitis vinifera
 98 Vitis vinifera
 99 Vitis vinifera
 100 Vitis vinifera

REFERENCE 1 Goes da Silva, F., Lim, H., Tandolino, A., Baek, J., Jones, K., Walker, M.A. and Cook, D.R.
 2 Transcriptional responses of Vitis vinifera to infection by the bacterial pathogen Xylella fastidiosa
 3 Unpublished
 4 Contact: Douglas Cook, PhD
 5 CAES Genome Facility
 6 UC Davis, Plant Pathology
 7 One Shields Ave, Davis, CA 95616, USA
 8 Tel: 530 754 6561
 9 Fax: 530 754 6617
 10 Email: drcoc@ucdavis.edu
 11 Seg primer: ACCGTACCGACATATGCC.
 12 Location/Qualifiers
 13 1..638
 14 /organism="Vitis vinifera"
 15 /mol_type="mRNA"
 16 /cultivar="Cabernet Sauvignon"
 17 /db_xref="taxon:29760"
 18 /clone="CA48N0001.11BF_G02"
 19 /sex="Hermaphrodite"
 20 /dev_stage="Berry stage I"
 21 /lab_host="DMSalpa"
 22 /clone_lib="Cabernet Sauvignon leaf-CA48N"
 23 /note="Organ: Berry; Vector: pDR, Site_1: Sfil; Site_2: Sfil; CA48N is a cDNA library of Cabernet Sauvignon leaves. The leaves were collected on July 25, 2001, in Napa Valley, California, and represent leaves in mid-season development. These leaves were asymptomatic and verified to be non-infected with the bacterial pathogen, Xylella fastidiosa, based on a diagnostic assay using PCR and Xylella-specific primer pairs. cDNA made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
 24 5'-AAGCAGTGTATCATACGACGAGTGCATTAAGCGCGG-3' and
 25 5'-ATTCTAGAGCGGAGCGGCGGACATG-3' (30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

BASE COUNT 133 a 172 c 189 g 144 t
 ORIGIN

Query Match 15.1%; Score 304; DB 14; Length 638;
 Best Local Similarity 69.2%; Pred. No. 2.6e-50;
 Matches 415; Conservative 0; Mismatches 185; Indels 0; Gaps 0;

QY 1351 TATTACGAGATGATATTCAGAGACACAGACAGTTCCTTATTCAGAGGATCC 1410
 Db 36 TCTTATCGGAGATGATATTCAGAGACACAGACAGTTCCTTATTCAGAGGATCC 95
 QY 1411 TCGCTAGATGATGCTTCTATTCAGATTCATCAACGTTCAAAACAGCTGATCAG 1470

Db 96 TCGAAGATGATCTACCAACCTGAGTTCGAATCAAGGATCCACAGATGAGATAGG 155
 QY 1471 TAGTGGGAAATCAAGAGCCCTAGATGATATGATCTGCTTACTCTCATATTAACG 1530
 Db 156 TGGTGGAAAGTCAAGGCGCGGTAAGAGATCCATTCGCGCATGATTAATCTGCGCG 215
 QY 1531 CTGTGGTAAAGAAATGATTTAGAGCTTCACTCTCCAGCCCACTTATCAATGAGCCCGTT 1590
 Db 216 CGGTGATGAGAGGATTTCTGAGGCTGACACCCAGCGGCCCACTCTCTGATGAGCCCGTT 275
 QY 1591 TGGCCATTAACACACAGATTCGTTGATGATGCTCTTTTCCGCGAGGAGACACAGCATAG 1650
 Db 276 TATCATATCAAGATTCACAGGATGATGATGATGATGATGATGATGATGATGATGATGAT 335
 QY 1651 TGAACATGAGGCGCGCTATCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 1710
 Db 336 TTAACATGAGGCGCGCTATCGATATGATGATGATGATGATGATGATGATGATGATGATGAT 395
 QY 1711 CTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1770
 Db 396 CCGACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 455
 QY 1771 GACTTGACCTTTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 1830
 Db 456 GCGTGGCGCGCTATCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 515
 QY 1831 CGGTATGATTTGAGCGGCGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1890
 Db 516 CAGCACCTTTGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGATGCGGAT 575
 QY 1891 ACGGCGTTGACTTATCTGAGAAAGTGAAGGCTTTTTCGAGATGATGATGATGATGATGAT 1950
 Db 576 ACCGATGATCTATCGAAGTGTGAGCTATCTCGAATGACACCCCTGACTG 635

RESULT 15
 BH581933 765 bp DNA linear GSS 15-DEC-2001
 LOCUS BOHRM099TR BOHR Brassica oleracea genomic clone BOHRM09, genomic survey sequence.
 BH581933
 ACCESSION BH581933.1 GI:17834390
 VERSION BH581933.1
 KEYWORDS GSS
 SOURCE Brassica oleracea
 ORGANISM Brassica oleracea

REFERENCE 1 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids
 2 Vitis vinifera
 3 Vitis vinifera
 4 Vitis vinifera
 5 Vitis vinifera
 6 Vitis vinifera
 7 Vitis vinifera
 8 Vitis vinifera
 9 Vitis vinifera
 10 Vitis vinifera
 11 Vitis vinifera
 12 Vitis vinifera
 13 Vitis vinifera
 14 Vitis vinifera
 15 Vitis vinifera
 16 Vitis vinifera
 17 Vitis vinifera
 18 Vitis vinifera
 19 Vitis vinifera
 20 Vitis vinifera
 21 Vitis vinifera
 22 Vitis vinifera
 23 Vitis vinifera
 24 Vitis vinifera
 25 Vitis vinifera
 26 Vitis vinifera
 27 Vitis vinifera
 28 Vitis vinifera
 29 Vitis vinifera
 30 Vitis vinifera
 31 Vitis vinifera
 32 Vitis vinifera
 33 Vitis vinifera
 34 Vitis vinifera
 35 Vitis vinifera
 36 Vitis vinifera
 37 Vitis vinifera
 38 Vitis vinifera
 39 Vitis vinifera
 40 Vitis vinifera
 41 Vitis vinifera
 42 Vitis vinifera
 43 Vitis vinifera
 44 Vitis vinifera
 45 Vitis vinifera
 46 Vitis vinifera
 47 Vitis vinifera
 48 Vitis vinifera
 49 Vitis vinifera
 50 Vitis vinifera
 51 Vitis vinifera
 52 Vitis vinifera
 53 Vitis vinifera
 54 Vitis vinifera
 55 Vitis vinifera
 56 Vitis vinifera
 57 Vitis vinifera
 58 Vitis vinifera
 59 Vitis vinifera
 60 Vitis vinifera
 61 Vitis vinifera
 62 Vitis vinifera
 63 Vitis vinifera
 64 Vitis vinifera
 65 Vitis vinifera
 66 Vitis vinifera
 67 Vitis vinifera
 68 Vitis vinifera
 69 Vitis vinifera
 70 Vitis vinifera
 71 Vitis vinifera
 72 Vitis vinifera
 73 Vitis vinifera
 74 Vitis vinifera
 75 Vitis vinifera
 76 Vitis vinifera
 77 Vitis vinifera
 78 Vitis vinifera
 79 Vitis vinifera
 80 Vitis vinifera
 81 Vitis vinifera
 82 Vitis vinifera
 83 Vitis vinifera
 84 Vitis vinifera
 85 Vitis vinifera
 86 Vitis vinifera
 87 Vitis vinifera
 88 Vitis vinifera
 89 Vitis vinifera
 90 Vitis vinifera
 91 Vitis vinifera
 92 Vitis vinifera
 93 Vitis vinifera
 94 Vitis vinifera
 95 Vitis vinifera
 96 Vitis vinifera
 97 Vitis vinifera
 98 Vitis vinifera
 99 Vitis vinifera
 100 Vitis vinifera

AUTHORS Town, C.D., Van Aken, S., Uteback, T., Koo, H. and Fraser, C.M.
 JOURNAL Whole genome shotgun sequencing of Brassica oleracea
 COMMENT Other_GSS: BOHRM099TR
 Contact: Chris Town
 TIGR
 9712 Medical Center Drive, Rockville, MD 20850, USA.
 Tel: 301-838-3523
 Fax: 301-838-0208
 Email: cdtown@tigr.org
 DNA is from a doubled haploid provided by Tom Osborn.
 Seg primer: TR
 Class: sheared ends.

FEATURES
 source
 1..765
 /organism="Brassica oleracea"
 /mol_type="Genomic DNA"
 /strain="TO1000D3"
 /db_xref="taxon:3712"
 /clone="BOHRM09"
 /clone_lib="BOHR"
 /note="Vector: PHOS1; Site_1: BstXI, 2-3 kb sheared genomic DNA inserted into PHOS1 using BstXI linkers"

BASE COUNT 197 a 167 c 192 g 209 t

Tue Jan 20 17:43:19 2004

us-10-022-025a-1.rst

Page 11

ORIGIN

Query Match 14.4%; Score 289.2; DB 28; Length 765;
Best Local Similarity 65.3%; Pred. No. 2.4e-47;
Matches 474; Conservative 0; Mismatches 243; Indels 9; Gaps 3;

```
QY 136 TTCCTTCTCTTGGCTTATCGGCATAGCGTAAAGCTCCTTAATCTTGC 195
Db 13 TTATTCAGTATGTTTCTCTGGTATAGAACCAATAGAAATCTGTGGCTTGC 72
QY 196 CTTTGTCCAAATGACAGCTTCTAAGCAAAACCACTTGCCTCTCCGCGT 255
Db 73 TCTTGATCAAAATGCAATATTCATTCAGAAATTTTCTCTCATCGCGT 132
QY 256 CACATCATCTGCTCGCATATCTCTTTATGACCTATCCGGTGGACCTGCTG 315
Db 133 AGCCATCACTGGTTTACCTTCACCATGTATTTGGTCTACCTGGAGCGGCTTG 192
QY 316 GGGAAATACCTCTTC---GGCGGTATATCCGGTTCATAC---AAACCGGAAACGT 369
Db 193 GGGAAATACCTTTTCACTCGCGGTTTACTTCGGCTGTAAACCAAAAACAGAAATCT 252
QY 370 TATTCCTGGTCCAAAGGCTTCCCTTGTGGTGAAGCATCTCACTCATGTCAAGCATCT 429
Db 253 CATTCCTGGTCCGAGAGGGTTTCCACTCGTGGGAAGCATGAGCTTAAGTCAAGCGCGT 312
QY 430 AGCTACCGGAGCATCGCTGATGACGTGAGAAATTCGAGCCAGAGGCTCATGGCTTT 489
Db 313 AGCTACCGGAGCATGACGAGCTGCTGCGATGAGAAACGCAAGCGGCTCATGGCGTT 372
QY 490 CAGCTTAGAGAGACTCCGCTGATGTCATCGTCAATCCGACGTAGCGAAAGATTTCT 549
Db 373 TAGCTTCGGTGTACCAAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 432
QY 550 GAATAGCCGGTTTTGTGATGACCGGTTAAAGATCGGTTACTCATGATGTTAA 609
Db 433 GAACGCTTGTGTTTGTGACGACGACTGTTGACGAAACGCTTATGTTGATGTTAA 492
QY 610 CAGAGCAATGTTTGTGACCAACGCTGTTTACTGGCGAACGCTTCCCGTATCGCTTC 669
Db 493 CCGAGCCATGGGATTTGCTCCAAGTGTGCTTACTGCGCAAGCTACGTCGTTAGGTTTC 552
QY 670 GAACCATCTCTTATGACAAACAAATGAGAGCCGAGCAACGACGATGATCTC 729
Db 553 GAACCTATCTTTTAAACCAAGCAATCAACGCTCGAGAGCAAGACGATGATGCTC 612
QY 730 AAGCCAGATGTTGATGTTCTTGAATAACAGATGATGATGATGATGATGATGATGATGAT 789
Db 613 GACTCGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 669
QY 790 GTTGTCTTAAACGGCTGCTTAAACGATGATGATGATGATGATGATGATGATGATGAT 849
Db 670 TTTGCTTAAACCGCTGCTTGTCTTAAACGATGATGATGATGATGATGATGATGATGAT 729
QY 850 GCTTGA 855
Db 730 ACTGGA 735
```

Search completed: January 16, 2004, 13:41:24
Job time : 4830 secs